

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

ANALYSIS OF SHIPHANDLERS' EYE-GAZE AND SIMULATION DATA FOR IMPROVEMENTS IN COVE-ITS SYSTEM

by

Salman Aleem

June 2017

Thesis Advisor:	Quinn Kennedy
Co-Advisor:	Perry McDowell
Second Reader:	Ruriko Yoshida

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** June 2017	**3. REPORT TYPE AND DATES COVERED** Master's thesis	
4. TITLE AND SUBTITLE ANALYSIS OF SHIPHANDLERS' EYE-GAZE AND SIMULATION DATA FOR IMPROVEMENTS IN COVE-ITS SYSTEM		**5. FUNDING NUMBERS** W7A95	
6. AUTHOR(S) Salman Aleem			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) Naval Research Program 699 Dyer Road, Halligan Hall, Rm 202. Monterey, CA 93943		**10. SPONSORING / MONITORING AGENCY REPORT NUMBER** NPS-17-N039-A	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Number: NPS.2017.0027-IR-EP7-A			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.		**12b. DISTRIBUTION CODE**	

13. ABSTRACT (maximum 200 words)

An intelligent tutoring system (ITS) is paired with U.S. Navy immersive virtual trainer for shiphandling called Conning Officer Virtual Environment (COVE) to monitor students' performance and provide spoken feedback. The feedback can be improved by incorporating the cognitive state of shiphandlers through analysis of their attention-allocation patterns from eye-tracking data. This thesis research contains a pilot study directed toward analyzing the eye-tracking data of expert and novice shiphandlers. We examined the relationship between a shiphandler's experience level, attention-allocation patterns, and performance during a simulated shiphandling exercise. Five novice and four expert shiphandlers from the U.S. Navy Surface Warfare Officers School participated in the study. Our analyses indicates expertise differences in shiphandling performance, general eye-tracking measures, scan transitions, as well as time distribution between different areas of interest. The experts' superior shiphandling performance was linked with having targeted and tight attention-allocation patterns that focused only on the relevant areas of interest. Novices' attention-allocation patterns were highly scattered and irregular. Results suggest that incorporating the "ideal" attention-allocation patterns of the experts into the ITS could improve its feedback to novice shiphandlers by telling novices where they should look and when. The study is based on a small sample size; therefore, further data collection should be performed to confirm the results.

14. SUBJECT TERMS shiphandling, eye tracking, intelligent tutor			**15. NUMBER OF PAGES** 137
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

ANALYSIS OF SHIPHANDLERS' EYE-GAZE AND SIMULATION DATA FOR IMPROVEMENTS IN COVE-ITS SYSTEM

Salman Aleem
Squadron Leader, Pakistan Air Force
BSc, University of Peshawar, 1999
M.S., Air University, Islamabad, 2015

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2017

Approved by: Quinn Kennedy
Thesis Advisor

Perry McDowell
Co-Advisor

Ruriko Yoshida
Second Reader

Patricia A. Jacobs
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

An intelligent tutoring system (ITS) is paired with U.S. Navy immersive virtual trainer for shiphandling called Conning Officer Virtual Environment (COVE) to monitor students' performance and provide spoken feedback. The feedback can be improved by incorporating the cognitive state of shiphandlers through analysis of their attention-allocation patterns from eye-tracking data. This thesis research contains a pilot study directed toward analyzing the eye-tracking data of expert and novice shiphandlers. We examined the relationship between a shiphandler's experience level, attention-allocation patterns, and performance during a simulated shiphandling exercise. Five novice and four expert shiphandlers from the U.S. Navy Surface Warfare Officers School participated in the study. Our analyses indicates expertise differences in shiphandling performance, general eye-tracking measures, scan transitions, as well as time distribution between different areas of interest. The experts' superior shiphandling performance was linked with having targeted and tight attention-allocation patterns that focused only on the relevant areas of interest. Novices' attention-allocation patterns were highly scattered and irregular. Results suggest that incorporating the "ideal" attention-allocation patterns of the experts into the ITS could improve its feedback to novice shiphandlers by telling novices where they should look and when. The study is based on a small sample size; therefore, further data collection should be performed to confirm the results.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION....1
A. LITERATURE REVIEW....2
1. Conning Officer Virtual Environment....2
2. Eye-Tracking Metrics and Uses....3
B. RESEARCH QUESTIONS....5

II. METHODOLOGY....7
A. PRELIMINARY STEPS PRIOR TO DATA COLLECTION....7
B. PARTICIPANTS....8
C. EQUIPMENT....9
D. SCENARIO....13
E. RECRUITMENT PROCEDURES....14
F. SURVEYS....15
G. PROCEDURES....15
H. EXPERIMENT PHASE CONCLUSION....16

III. DATA PREPARATION....17
A. EYE-TRACKING DATA....17
1. Eye-Gaze Metrics....24
2. Raw Data Export....25
B. SIMULATOR LOGS....27
C. VOICE LOGS....29
D. DATA SYNCHRONIZATION....31

IV. ANALYSES AND RESULTS....33
A. EXPERTISE DIFFERENCES IN SHIPHANDLING PERFORMANCE....33
1. Ship's Path Visualization....34
2. Time Taken by Participants....35
B. EYE-TRACKING FIXATION METRICS....37
1. General Attention Allocation Metrics....37
2. Eye-Scan Graphs....46
C. PERFORMANCE ANALYSIS FROM SIMULATION DATA....54
1. Measures of Performance....54
2. Criteria and Evaluation....55
3. Analytical Hierarchy Process....55
4. Expert vs. Novice Performance....56

5. Novices' Performance Evaluation 57
D. POST TASK SURVEYS....59

V. CONCLUSION AND DISCUSSION 63
A. SUMMARY OF RESULTS 63
B. IMPLICATIONS OF RESULTS....63
1. Results for Classroom Teaching....63
2. Update the Expert Model in ITS 64
3. Live Multi-student Performance Display for Instructors 64
4. Debrief System 65
C. LIMITATIONS....65
1. Small Sample Size 65
2. Eye-Tracking Data Collection Problems 66
3. Data Loss in Simulation and Voice Logs 69
D. RECOMMENDATIONS ON FUTURE WORK 69

APPENDIX A. DEMOGRAPHICS SURVEY FORM 71

APPENDIX B. POST-TASK SURVEY FORM....73

APPENDIX C. EYE-TRACKING METRICS SEGMENTS- 1 AND 2 75

APPENDIX D. EYE SCAN AND AOI TIME DISTRIBUTION GRAPHS 81

LIST OF REFERENCES....109

INITIAL DISTRIBUTION LIST 113

LIST OF FIGURES

Figure 1. Typical Setup of COVE-3 Simulator 10

Figure 2. View of COVE-3 from Student's Position. 11

Figure 3. Tobii Pro Glasses 2 Used for Eye Tracking 12

Figure 4. Approach Channel and Pier Setting of Port Mina Salman, Bahrain for the Simulation Exercise in COVE-3 14

Figure 5. Tobii Pro Lab Software — Project Overview for Our Study's Data 17

Figure 6. Recorded Video Display in Tobii Pro Lab with Superimposed Gaze Position 18

Figure 7. Areas of Interest Defined to Map Fixation Data 20

Figure 8. Marking the Fixation Information from Video on the AOIs Chart 23

Figure 9. Events Definition on Top and as Marked in a Recording at the Bottom 23

Figure 10. Simulator Log File Objects and Attributes 28

Figure 11. Voice Log Files Format 29

Figure 12. Ships' Path in Simulation Exercise 34

Figure 13. Participants' Paths during Segments 3 and 4 35

Figure 14. Time Taken by Each Participant to Complete the Exercise (Segmented) 35

Figure 15. Time Taken to Complete the Exercise (Segmented and Grouped) 36

Figure 16. AOI Fixation Count Percentages in Segment 3 38

Figure 17. AOI Fixation Duration Percentages in Segment 3 39

Figure 18. AOI Average Fixation Duration in Segment 3 41

Figure 19. AOI Fixation Count Percentages in Segment 4 42

Figure 20. AOI Fixation Duration Percentages in Segment 4 44

Figure 21. AOI Average Fixation Duration in Segment 4 46

Figure 22. Eye-Scan Network and Timeline Graphs of P-7-7 in Segment 3 49
Figure 23. Eye-Scan Network and Timeline Graphs of I-7-4 in Segment 3 50
Figure 24. Eye-Scan Network and Timeline Graphs of P-7-7 in Segment 4 52
Figure 25. Eye-Scan Network and Timeline Graphs of I-7-4 in Segment 4 53
Figure 26. Parameters Used for Evaluating Shiphandler's Performance 55
Figure 27. Comparison of Expert and Novice Performance 57
Figure 28. Performance Evaluation of Novices 58
Figure 29. Impression of COVE-ITS Live Performance Monitoring System 65
Figure 30. Track Status Tool Guidelines from Tobii Pro Glasses 2 User Manual.....67
Figure 31. Screenshot of Eyes as Recognized from Glasses Controller Software.....68
Figure 32. AOI Fixation Count Percentages in Segment 1 75
Figure 33. AOI Fixation Duration Percentages in Segment 1.....76
Figure 34. Mann-Whitney U-test Result for Seg-1 Fixation Duration Percentage 76
Figure 35. Average Fixation Duration in AOIs in Segment 1.....77
Figure 36. AOI Fixation Count Percentages in Segment 2 78
Figure 37. AOI Fixation Duration Percentages in Segment 2.....79
Figure 38. Average Fixation Duration in AOIs in Segment 2.....80
Figure 39. I-6-2 Eye Scan Network Graph in Segment 1 81
Figure 40. I-6-2 Eye Scan Timeline Graph in Segment 1 81
Figure 41. I-6-3 Eye Scan Network Graph in Segment 1 82
Figure 42. I-6-3 Eye Scan Timeline Graph in Segment 1 82
Figure 43. I-7-4 Eye Scan Network Graph in Segment 1 83
Figure 44. I-7-4 Eye Scan Timeline Graph in Segment 1 83
Figure 45. I-7-5 Eye Scan Network Graph in Segment 1 84
Figure 46. I-7-5 Eye Scan Timeline Graph in Segment 1 84

Figure 47. I-8-8 Eye Scan Network Graph in Segment 1 85
Figure 48. I-8-8 Eye Scan Timeline Graph in Segment 1 85
Figure 49. III-8-10 Eye Scan Network Graph in Segment 1 86
Figure 50. III-8-10 Eye Scan Timeline Graph in Segment 1 86
Figure 51. P-7-7 Eye Scan Network Graph in Segment 1 87
Figure 52. P-7-7 Eye Scan Timeline Graph in Segment 1 87
Figure 53. P-8-9 Eye Scan Network Graph in Segment 1 88
Figure 54. P-8-9 Eye Scan Timeline Graph in Segment 1 88
Figure 55. I-6-2 Eye Scan Network Graph in Segment 2 89
Figure 56. I-6-2 Eye Scan Timeline Graph in Segment 2 89
Figure 57. I-6-3 Eye Scan Network Graph in Segment 2 90
Figure 58. I-6-3 Eye Scan Timeline Graph in Segment 2 90
Figure 59. I-7-4 Eye Scan Network Graph in Segment 2 91
Figure 60. I-7-4 Eye Scan Timeline Graph in Segment 2 91
Figure 61. I-7-5 Eye Scan Network Graph in Segment 2 92
Figure 62. I-7-5 Eye Scan Timeline Graph in Segment 2 92
Figure 63. I-8-8 Eye Scan Network Graph in Segment 2 93
Figure 64. I-8-8 Eye Scan Timeline Graph in Segment 2 93
Figure 65. III-8-10 Eye Scan Network Graph in Segment 2 94
Figure 66. III-8-10 Eye Scan Timeline Graph in Segment 2 94
Figure 67. P-7-7 Eye Scan Network Graph in Segment 2 95
Figure 68. P-7-7 Eye Scan Timeline Graph in Segment 2 95
Figure 69. P-8-9 Eye Scan Network Graph in Segment 2 96
Figure 70. P-8-9 Eye Scan Timeline Graph in Segment 2 96
Figure 71. I-6-2 Eye Scan Network Graph in Segment 3 97

Figure 72. I-6-2 Eye Scan Timeline Graph in Segment 3 97
Figure 73. I-6-3 Eye Scan Network Graph in Segment 3 98
Figure 74. I-6-3 Eye Scan Timeline Graph in Segment 3 98
Figure 75. I-7-5 Eye Scan Network Graph in Segment 3 99
Figure 76. I-7-5 Eye Scan Timeline Graph in Segment 3 99
Figure 77. I-8-8 Eye Scan Network Graph in Segment 3 100
Figure 78. I-8-8 Eye Scan Timeline Graph in Segment 3 100
Figure 79. III-8-10 Eye Scan Network Graph in Segment 3 101
Figure 80. III-8-10 Eye Scan Timeline Graph in Segment 3 101
Figure 81. P-8-9 Eye Scan Network Graph in Segment 3 102
Figure 82. P-8-9 Eye Scan Timeline Graph in Segment 3 102
Figure 83. I-6-2 Eye Scan Network Graph in Segment 4 103
Figure 84. I-6-2 Eye Scan Timeline Graph in Segment 4 103
Figure 85. I-6-3 Eye Scan Network Graph in Segment 4 104
Figure 86. I-6-3 Eye Scan Timeline Graph in Segment 4 104
Figure 87. I-7-5 Eye Scan Network Graph in Segment 4 105
Figure 88. I-7-5 Eye Scan Timeline Graph in Segment 4 105
Figure 89. I-8-8 Eye Scan Network Graph in Segment 4 106
Figure 90. I-8-8 Eye Scan Timeline Graph in Segment 4 106
Figure 91. III-8-10 Eye Scan Network Graph in Segment 4 107
Figure 92. III-8-10 Eye Scan Timeline Graph in Segment 4 107
Figure 93. P-8-9 Eye Scan Network Graph in Segment 4 108
Figure 94. P-8-9 Eye Scan Timeline Graph in Segment 4 108

LIST OF TABLES

Table 1. Demographics of Participants 9

Table 2. Specification of Tobii Pro Glasses 2 12

Table 3. Parameters of Tobii I-VT (Attention) Filter (Courtesy or Tobii AB). Adapted from Tobii Pro Lab software. 19

Table 4. Description of AOIs Used in Eye Tracking 20

Table 5. Total Fixation Duration (Seconds) for Some AOIs in Segment 4 (Sample) 25

Table 6. Total Fixation Count for Some AOIs in Segment 4 (Sample) 25

Table 7. Example of Exported Raw Data Information 26

Table 8. Extracted Data Example From Simulator Log Files in CSV Format 29

Table 9. Extracted Table Format of Voice Commands 30

Table 10. T-test and MW U-test Result for Seg-3 Fixation Count Percentage 39

Table 11. T-test and MW U-test Result for Seg-3 Total Fixation Duration Percentage 40

Table 12. T-test and MW U-test Result for Seg-3 Average Fixation Duration 41

Table 13. T-test and MW U-test Result for Seg-4 Fixation Count Percentage 43

Table 14. T-test and MW U-test Result for Seg-4 Total Fixation Duration Percentage 45

Table 15. T-test and MW U-test Result for Seg-4 Average Fixation Duration 46

Table 16. Weights Matrix for AHP 55

Table 17. Self-Evaluation, Strategies and Visual Cues Used by Participants from Start Point to Approaching the Berth 60

Table 18. Self-Evaluation, Strategies and Visual Cues Used by Participants for Getting the Ship Lined up with the Pier 61

Table 19. Self-Evaluation, Strategies and Visual Cues Used by Participants for Bringing Ship Closer to the Pier 62

Table 20. Mann-Whitney U-test Result for Seg-1 Fixation Count Percentage 75

Table 21. Mann-Whitney U-test Result for Seg-1 Avg Fixation Duration Percentage 77

Table 22. Mann-Whitney U-test Result for Seg-2 Fixation Count Percentage 78

Table 23. Mann-Whitney U-test Result for Seg-2 Fixation Duration Percentage 79

Table 24. Mann-Whitney U-test Result for Seg-2 Average Fixation Duration Percentage 80

LIST OF ACRONYMS AND ABBREVIATIONS

ADOC	Advanced Division Officers Course
AHP	analytical hierarchical process
AOI	area of interest
COVE	Conning Officer Virtual Environment
CSV	Comma Separated Value
ETG	eye tracking glasses
HITL	human in the loop
IRB	Institutional Review Board
ITS	intelligent tutoring system
MC	major Command
MOP	measure of performance
OTW	out the window
PCO/XO	Prospective Commanding Officer/Executive Officer
SWO	surface warfare officer
SWOS	Surface Warfare Officers School
TFH	total flight hours
XML	Extensible Markup Language

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Background. The Conning Officer Virtual Environment (COVE) is an immersive virtual simulation system to train shiphandling students in U.S. Navy. An intelligent tutoring system (ITS) developed by Stanford University is paired with COVE to monitor students' performance and provide spoken feedback. One of the limitations in ITS feedback is that it does not include any information regarding "why" an individual student made an error (i.e., their cognitive state). A 2007 *Human Factors* article by Sarter et al. showed that attention-allocation patterns as detected through eye-tracking can be used to predict an individual's level of task experience as well as underlying cognitive strategies in areas such as overland navigation, flight control, and driving. We are of the opinion that the ITS feedback can be improved by incorporating the cognitive state of shiphandlers through analysis of their attention-allocation patterns from eye-tracking data.

Purpose. The purpose of this thesis is to conduct a pilot study to begin to determine whether the addition of eye tracking to the ITS would provide more targeted training. The pilot study collected eye-tracking data of shiphandlers during a COVE scenario and analyzed their Attention-allocation patterns along with performance data.

Methods. Five novice and four expert shiphandlers from the U.S. Navy Surface Warfare Officers School participated in the study. The scenario selected for the study in consultation with the sponsor was a capstone exercise for a novice group that entails mooring a DDG-51 at Mina Salman, Bahrain as a conning officer. The eye-tracking data during the simulated exercise was collected using wearable eye tracking glasses (ETG) manufactured by Tobii AB, Sweden.

Analysis/Results. There are three principal data sources in our study besides surveys: eye-tracking data, simulator data, and voice logs. Eye-tracking data was processed through manual mapping of fixation information on Areas of Interest (AOI) chart and the extracted information was time synchronized with simulation data and voice logs. As per sponsor's input, the transit was divided into four segments and the last two

segments (3rd and 4th) were analyzed in more details as they are most error prone. The third segment starts when the shiphandler's ship crosses the last pair of buoys in the channel while the fourth segment starts when the stern of the shiphandler's ship clears the already berthed ship at the edge of pier and finishes on successful landing.

Our analyses indicate expertise differences in shiphandling performance, general eye tracking measures, scan transitions, as well as time distribution between different areas of interest. There was less variability in experts' completion time of different exercise segments while among novices there was much variability. The positional information of ships when plotted on the map indicated experts' routes in spatial proximity. On the other hand, novice routes were disparate and some even could not successfully moor their ships. The fixation metrics also indicated experts' looking at desirable areas of interest (AOIs) more than the novices. The fixations when viewed as a network of transitions between different AOIs indicated experts fixating more at some key AOIs during a particular situation while among novices there were a lot of fixations at a greater number of AOIs. The time distribution of fixation also reinforced the same effect as network of transitions.

In summary, the experts' superior shiphandling performance was linked with having targeted and tight Attention-allocation patterns that focused only on the relevant areas of interest. Novices' Attention-allocation patterns were highly scattered and irregular. Results suggest that incorporating the "ideal" Attention-allocation patterns of experts into ITS could improve its feedback to novice shiphandlers by telling novices where they should look at and when.

Conclusion and Future Work. This work can be applied to any military task where use of visual cues is necessary and more experience results in efficient scan patterns. The information gained from this study can be used in teaching optimal attention allocation techniques to trainee shiphandlers during various phases of mooring exercise. The expert model of ITS, which is a cognitive model that represents an expert's performance for various shiphandling tasks has incorporated visual cues which were discovered through interviews with subject matter experts and head tracking. However, these are coarse techniques as attention allocation becomes automatic after mastering a

skill and may not be fully conveyed by the experts. The head tracking also captures only the general direction of focus and not the exact area of interest being fixated by eyes. In contrast, the information gained through eye tracking in this study can provide accurate quantitative data on visual cues used by experts and hence can be used to update the expert model of ITS which may result in better feedback to the students. The algorithm developed in this study to analyze the novice performance based upon experts' simulation data weighed as per analytical hierarchical process (AHP) can be a developed into a tracking tool for live performance monitoring of the students. The data from eye-tracking hardware can also be used in conjunction with performance evaluation to develop a de-brief system where instructors can show the students their Attention-allocation patterns during various phases of exercise along with their performance and can coach them on their cognitive skills and ideal cognitive behavior. The results obtained in this pilot study may be complemented with further data collection and analyses.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank Dr. Quinn Kennedy and Mr. Perry McDowell for giving me the opportunity to work with them and guiding me all the way toward its accomplishment. Dr. Rudy Yoshida's statistical insights were key to my analyses and also key for successful completion. Research Associate Rabia Khan's work in extracting fixation information was a vital step that she did very well. Together, all four were a wonderful team who helped me greatly in reaching my goal.

I am thankful to people at SWOS, CAPT Rich Callas, CAPT Bud Weeks, CAPT Ed Rogers, Richard Kulesh and LT Bob Burke for their support during the study and also during our visits for data collection. I owe gratitude to Stanley Peters and Jia Huang at CSLI, Stanford University, for giving an orientation of COVE-ITS system and their help in understanding the simulation data format. I would also like to thank the U.S. Navy officers at SWOS who volunteered for this study and the Naval Research Program (NRP) for its funding, which made it possible.

I would like to thank NPS and the operations research faculty for the great learning experience they provided me in past 20 months and to the U.S. IMET program for funding my study. I am greatly indebted to my country, Pakistan, and the Pakistan Air Force, for providing me the educational opportunity in the USA. I would also like to thank my fellow student and long-time friend, Faisal, for his support and guidance.

I would like to dedicate my work to my late father who always encouraged me to keep learning and doing work to benefit my country. Words cannot express how grateful I am to my mother for all the sacrifices she made on my behalf. Her prayers for me were what sustained me thus far. I am grateful to my dear brother and sisters for their prayers and support. In the end, I would like to do a very special thanks to my beloved wife, Maham, for her support, love, kindness, and sacrifice. She spared me time by taking care of the house and our four children. She really encouraged me to do this work better and better and provided support to accomplish it.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The Conning Officer Virtual Environment (COVE) is a virtual simulation system to train shiphandling students. It has been used by the U.S. Navy to train surface warfare officers (SWOs) in acquiring shiphandling skills at a number of locations within the United States for over 15 years. It is used for very basic to advanced levels of training and provides realistic graphics to practice shiphandling skills. The system, however, suffers from being human-effort-intensive as it requires one-on-one tutoring (i.e., one instructor is required per student for each training exercise). To overcome this limitation, an intelligent tutoring system (ITS) was developed by Stanford University and was added to COVE around 2010 (Kirschenbaum et al., 2010). The ITS monitors the student's performance during training and provides spoken feedback on errors he/she is making and guidance to correct those errors. A limitation of this feedback in case of an error is that it is based on the assumption that the cause of error is the most statistically common one. There can be a number of reasons behind an error, however, so every error may not be the statistically most common. The ITS also does not include any information regarding "why" an individual student made an error (i.e., their cognitive state).

Attention-allocation patterns as detected through eye tracking have been shown to predict an individual's level of task experience as well as underlying cognitive strategies in areas such as overland navigation, flight control, and driving (Kirby et al., 2014; Falkmer, 2001). The eye tracking devices enable us to record data on where, when and for how long a trainee looks at relevant and irrelevant pieces of information during a training scenario. Previous studies have indicated that attention allocation data can provide valuable training information to instructors, for example, that the trainee is off-course because they are focusing on the wrong visual cues during critical times of decision making (Sarter et al., 2007; Schriver et al., 2008; Sullivan, et al., 2011; Yang et al., 2013). Thus, eye-tracking data has the potential to provide more indications as to why an individual student made an error.

The purpose of this thesis is to conduct a pilot study to begin to determine whether the addition of eye tracking to ITS would provide more targeted training. The

pilot study examined Attention-allocation patterns and performance data of shiphandlers during a COVE scenario who ranged in their experience levels from 2 to 30 years. This study will also extend previous findings by examining Attention-allocation patterns in a new population (shiphandlers) and incorporating novel statistical and visualization methods in the eye-tracking data.

A. LITERATURE REVIEW

Simulation is a useful, safe and cost-effective method for education, training, planning and analysis (Kincaid, Hamilton, Tarr, & Sangani, 2003; Keller-McNulty et al., 2006). In the military, simulations are generally categorized into live, virtual and constructive domains owing to what is being simulated. Live simulations involve real people and platforms but the environment and some of the weapons onboard platforms may be simulated (Davis, 1995). Virtual simulations involve real people but simulated platforms, while constructive simulations involve both simulated people and platforms (Davis, 1995).

Virtual simulations are also called human-in-the-loop (HITL) simulations due to the involvement of human operators (Davis, 1995). One of the main uses of this kind of simulation is in acquiring skills to operate the platform that is being simulated (e.g., flight simulators, ship driving simulators, vehicle driving simulators, weapons simulators). The objectives in using virtual simulations for training are to reduce the cost of training, provide a safer environment and have more availability of equipment for training (Iyengar et al., 1999).

1. Conning Officer Virtual Environment

COVE uses an ITS system which comprises an expert cognitive model and an intelligent tutor (Wong, Kirschenbaum, & Peters 2010). The expert model is a Java-based implementation of Adaptive Control of Thought—Rational (ACT-R) cognitive model architecture (Anderson et al., 2004), which outputs the expected expert's actions under current conditions to ITS. Wong et al. (2010) have described that a shiphandling task may be performed in many correct ways and different experts may use similar or different strategies to accomplish a particular task. Therefore, the expert model has multiple action

paths and is based upon flawless performance and perceptual heuristics. The visual cues used by experts are extracted from the objects in the field of view of a head-mounted display (COVE-1) or the center screen of multiple displays (COVE-3) for input into the expert model. The scanning patterns between different areas during an exercise are incorporated into the model through experts' opinions and head tracking data.

The student actions while performing a simulation on COVE-ITS are compared with the expert model in ITS. This expert model informs the intelligent tutor about most probable expert actions under the current situation and their rationale. The intelligent tutor then integrates student and expert actions, carries out its measurement and gives feedback to the student (Wong et al., 2010).

2. Eye-Tracking Metrics and Uses

There are a number of metrics used in eye tracking research. Holmqvist et al. (2011) have described fixation as the most widely used and reported event. Fixation is defined as a state in which the eye remains still over a period of time, and it can last from some milliseconds to several seconds. The other important metric is saccade, which is rapid eye movement between fixations that lasts typically 30–80 milliseconds. The human eye is considered to be blind during a saccade. Eye tracking has been used in a variety of studies including flying (Kirby et al., 2014) and car driving (Falkmer, 2001) and is described as a research tool that can be used in cognitive and usability analyses (Holmqvist et al., 2011).

Sullivan, Yang, Day, and Kennedy (2011) investigated the influence of expertise among helicopter pilots on visual scan patterns. They found that expertise level can be predicted by the general gaze parameters as well as visual scan techniques. More experienced pilots had shorter fixation durations and a faster scan rate. The more experienced pilots also spent more time looking out-the-window (OTW) than novices and were also found to confirm the visited position on the map apart from looking ahead on the map. Sullivan et al. (2011) deduced that eye trackers can be used to indicate the aspects of a trainee's cognitive state that can be used as input to tutoring systems.

Yang, Kennedy, Sullivan, and Fricker (2013) assessed the effect of expertise on helicopter pilots' gaze measurements and navigation accuracy in overland navigation on easy and difficult routes. Expertise was measured as total flight hours (TFH) for individual pilots. They found that while TFH was not related with error in any route section, experts spent less time scanning OTW and had shorter OTW dwell on easy routes. On difficult routes, experts were found to slow down their OTW scan and had shorter dwell. Thus, during these difficult routes, the experts' Attention-allocation patterns were similar to what is typically seen among less experienced pilots. These results indicate that within-person changes in Attention-allocation patterns can indicate when experts have shifted from a relatively easy task to a more challenging task.

The above studies used typical methods to examine eye gaze differences. In a series of studies, Horiguchi and collaborators used two novel techniques: sequence alignment graphs and graph clustering (Horiguchi et al., 2015; Horiguchi, Suzuki, Sawaragi, Nakanishi, & Takimoto, 2016). For sequence alignment, they used the CLUSTAL algorithm (Thompson, Higgins, & Gibson, 1994), which consists of aligning the pair of sequences to first calculate the distance matrix and then a guide tree followed by branch order progressive alignment in the guide tree. For graph clustering, they used the Markov Cluster Algorithm (van Dongen, 2001, 2008), which uses random walks to identify dense regions where graph nodes have many connections, thus forming a cluster in the graph.

In Horiguchi et al. (2015), sequence alignment graphs were used to analyze eye gaze patterns of high-speed train drivers in Japan. They found that experienced drivers shared more common Areas of Interest (AOI) than novices. They were less reliant on the speedometer to judge speed than novices, suggesting they may have depended more on optical flow motion looking outside through the windshield. The sequence alignment graphs revealed that the experts were more consistent in their timing to focus on visual cues at important times like when preparing to stop at a station. In a later study that also examined high-speed train drivers, Horiguchi et al. (2016) used the Markov Cluster Algorithm to structure the drivers' eye-scan transitions into clusters to find the patterns of frequent occurrence. One strong pattern observed among all drivers was to look ahead

after visiting any other area of interest. Experts were able to follow this pattern more consistently than novices, however.

Forsman, A. Dahlman, J. Dahlman, Falkmer, and Lee (2012) used eye tracking during high-speed navigation at sea to detect expertise differences in navigational gaze behavior. Novices were found to look more at objects closer to the boat while experts looked more at things far from the boat. Novice boat drivers were more focused on electronic displays, especially during high-speed conditions, while the experts were focused mostly outside the boat. Experts also used more paper charts to confirm surrounding environment than novices. The Forsman group's study shows use of eye tracking to find expertise difference among shiphandlers, which is also the purpose of our study. They used eye tracking in actual boat driving conditions, while our study was conducted in simulator. They took eye tracking metrics looking at inside gadgets or four sides outside (i.e., ahead, port, starboard and astern) during high-speed conditions, while our study is based upon a mooring scenario for a destroyer class ship.

B. RESEARCH QUESTIONS

Because a pilot study was conducted, no specific hypotheses were tested. Below are the main and subsidiary research questions.

Main Questions

- What is the relationship between shiphandler's experience level and scan techniques?
- Is the expert's attention-allocation pattern associated with better performance?
- When errors occur, does eye-tracking data provide insights into how and when the trainee's attention-allocation pattern deviated from the experts' Attention-allocation pattern?

Subsidiary Question (Subject to Availability of Time)

- How can the ideal attention-allocation pattern be integrated into ITS to provide improved feedback to the shiphandlers during COVE scenarios?

THIS PAGE INTENTIONALLY LEFT BLANK

II. METHODOLOGY

This project entailed multiple preliminary steps prior to actual data collection, including meetings with the ITS development team and the sponsor as well as the IRB application. These steps are described below. Next, descriptions of the participants, equipment, and procedures for the actual study are provided.

A. PRELIMINARY STEPS PRIOR TO DATA COLLECTION

These initial steps started with a literature review of the published articles on ITS and a meeting with the ITS development team at Stanford University to get experience using the COVE-ITS systems installed there. The Stanford ITS team also sent us the simulation and voice logs of some of the novices and experts shiphandling in our required scenario.

Another preliminary step was a trip to Surface Warfare Officer School (SWOS), Newport, Rhode Island, to meet the sponsor and gain firsthand knowledge about the problem as well as to discuss the study design and logistics of conducting the experiment at SWOS. The instructional staff demonstrated operation of the COVE simulator and conducted the required training scenario from Advanced Division Officers Course (ADOC) curriculum. We also practiced using the eye tracking glasses during that exercise with members of the study team. My co-advisor also liaised with school command on available dates for running the experiment. The aim was to have an ADOC course nearing graduation, when they have completed our required scenario in their curriculum. SWOS informed the research team of the best time to visit to recruit volunteer students and instructors and conduct the study. After necessary coordination on schedule and availability of COVE at SWOS for experiments, Institutional Review Board (IRB) approval had to be sought from the Naval Postgraduate School for using human subjects in the study.

After the first trip, the research team asked SWOS to formally send a sponsor letter in order to initiate the IRB process. A consent form to participate in research was prepared for potential volunteers; a pre-task survey was designed to collect the

demographics of participants; and a post-task survey was made to collect information about participants' own views about their performance, their strategies during different phases of simulation exercise and the visual cues they used. The scenario was finalized and a checklist was prepared for smooth and efficient conduct of the experiment and collection of data. All were submitted along with IRB application for review by the board and approval was obtained before beginning the study at SWOS.

B. PARTICIPANTS

The study participants were shiphandlers recruited from SWOS. The aim of the study was to have at least two distinct groups (i.e., novices and experts). SWOS conducts courses year-round for officers of different levels of commissioned service. The first time SWOs attend classes in Newport is for the Advanced Division Officers Course (ADOC). The ADOC students generally have 2–4 years of commissioned service and have completed one tour at sea. This community formed the novice group of our study. For experts, we mainly focused on recruiting the instructors stationed at SWOS to volunteer for the study. Students attending the Prospective Commanding Officers/Executive Officers (PCO/XO) Course having 15–19 years' service or the Major Command (MC) course with current experience on a DDG-51 destroyer were also considered for the expert group.

Eleven SWOs volunteered for the study, ten males and one female. One student was excluded because he had to wear eyeglasses and the eye tracker cannot be used with eyeglasses. The female student faced continuous calibration failure on the eye tracking glasses. Hence, we had nine participants who completed the study: five ADOC students, one ADOC instructor, two officers attending the PCO/XO course, and one officer attending the MC course. The instructor did two experimental runs on slightly different systems. The MC course participant was a very experienced shiphandler, but he had not had DDG-51 experience since early in his career. After analyzing his eye tracking and simulation data, his data had to be excluded from the study. Thus, results are based on five novice and four expert recordings; two of the expert recordings were performed by the same individual using slightly different approaches.

Table 1 outlines the demographic characteristics of the participants. The red shaded rows indicate the novice participants while the green shade specifies experts. All participants were accustomed to using the shiphandling simulator and believed the simulation to be an effective learning tool for shiphandling.

Table 1. Demographics of Participants

ID	Age	Years' Service	Highest Tour Completed	Months Aboard (DDG-51)	Months U/W Watch on Bridge (DDG-51)	Sea and Anchor Transits as Conn Off/OOD/XO/ CO (DDG-51)
I-6-2	24	2	1st DIVO	24	12	25
I-6-3	25	2	1st DIVO	30	14	10
I-7-4	25	2.25	1st DIVO	NA	NA	NA
I-7-5	28	6.5	1st DIVO	33	32	20
I-8-8	24	3	1st DIVO	32	32	70
III-8-10	29	6	2nd DIVO	54	35	15
P-7-7	39	15	DH	40	30	OOD/CON
P-8-9	45	21	CO	60	15	25

C. EQUIPMENT

a. Conning Officer Virtual Environment

The study was conducted at SWOS, Newport, with a COVE simulation system, which is a HITL trainer for shiphandling developed by a civilian company, CSRA. It has the capability to simulate all types of ships in the U.S. Navy. COVE has two variants based upon the difference in display technology used.

We used the COVE-3 simulator for our study, which uses three 85-inch displays for front display and a projector for stern view. The student stands facing the large center screen. A 20-inch instrument panel screen and a secondary view screen with joystick control are placed in front of him. The typical setup of COVE-3 systems installed at SWOS, RI is shown in Figures 1 and 2.

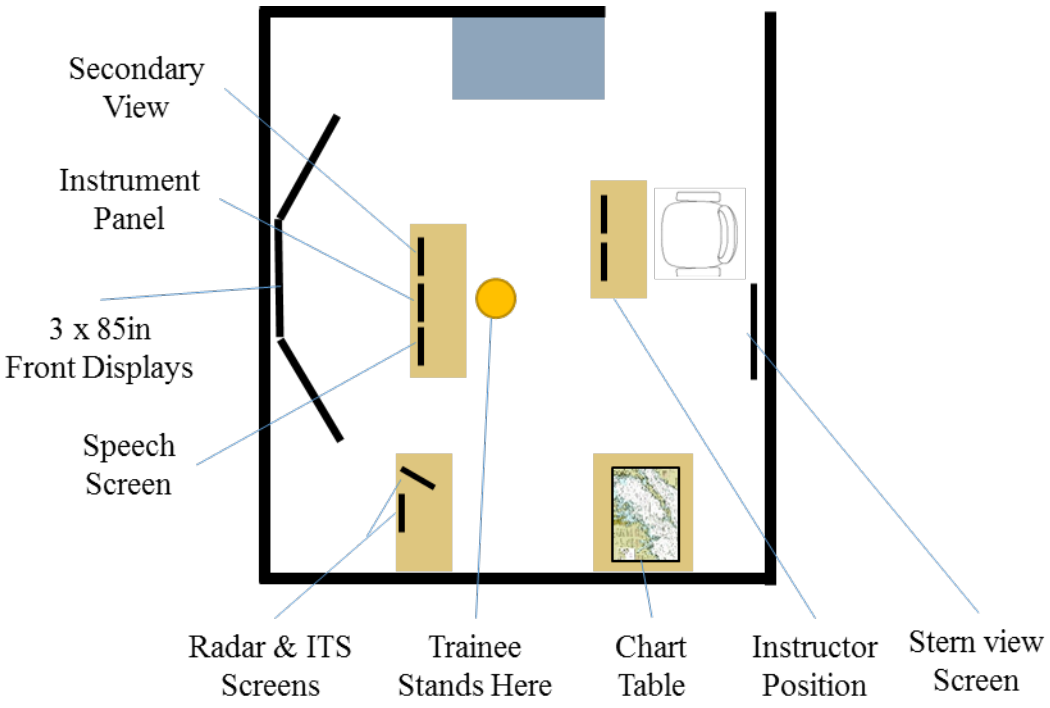


Figure 1. Typical Setup of COVE-3 Simulator

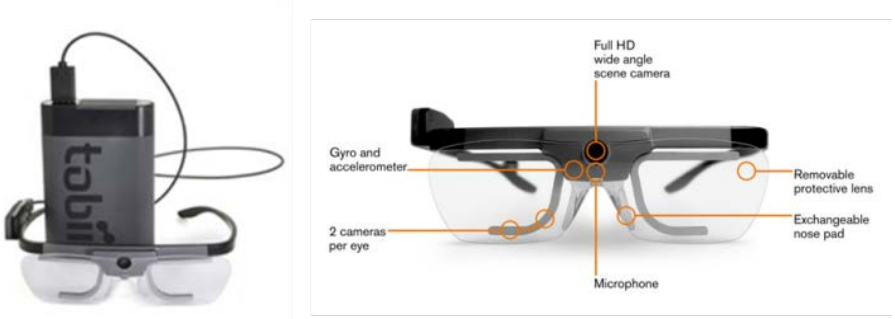


Figure 2. View of COVE-3 from Student's Position.

There are a number of COVE-3 systems installed in SWOS. One of them is in their test lab. The only difference between the test lab system and other systems is the size of the front screens (test lab front screens are 55-inch instead of 85-inch as in student labs). In our study, one instructor and two students performed simulation exercise in the test lab while the other six recordings were made in the COVE-3 lab. The instructors provide oversight to the simulation exercise by their physical presence and by watching the student's conduct at the Instructor Operator Station.

b. Tobii Pro Glasses-2

The eye-tracking data was collected using Tobii Pro Glasses 2, manufactured by Tobii AB, Sweden. These eye tracking glasses (ETG) are currently the latest release by Tobii in wearable technology. They are lightweight and comfortable to wear (see Figure 3). Their specifications are listed in Table 2.



Adapted from https://www.tobiipro.com/product-listing/tobii-pro-glasses-2/. Photo Courtesy of Tobii AB.

Figure 3. Tobii Pro Glasses 2 Used for Eye Tracking

The hardware consists of two units, the head unit and the recorder. The head unit consists of the eye tracking glasses with infrared illuminators and sensors installed in them along with an HD camera for recording front view and audio. The head unit connects to a portable recorder unit via an HDMI cable. The recorder operates on rechargeable batteries and is controlled through PC-based software through an active Wi-Fi connection.

Table 2. Specification of Tobii Pro Glasses 2.

Eye tracking technique	Corneal reflection, dark pupil
Sampling Rate	50 Hz/100 Hz
Field of View	82 deg horizontal, 52 deg vertical
Gaze Tracking Accuracy	0.5 deg
Gaze tracking range	>160 deg horizontal, 70 deg vertical
Scene camera resolution	1920 x 1080 at 25 fps
Frame Dimension (W x H)	179mm x 57mm
Weight	45 g
Interchangeable nose piece	Yes (2)

Adapted from Tobii Pro Glasses 2 user manual (2016)

The recording is performed on an SD card inserted in the recorder's slot. The ETG controller software runs on a Windows-based tablet/laptop/PC. After establishing Wi-Fi connection between tablet/laptop/PC and ETG recorder, battery health and memory card capacity can be checked on the software interface. The controller software also shows the live picture of the camera mounted on the head unit, and a circle superimposed on the video indicates the gaze position of the eyes. The calibration and recording on/off control also are performed via Tobii controller software.

D. SCENARIO

The scenario selected for the study in consultation with the sponsor was an ADOC course capstone assessment exercise that entails mooring a DDG-51-class ship at Mina Salman, Bahrain as a conning officer (Figure 4). This exercise demands skillful navigation and control of the ship as the channel is narrow and the pier is usually busy with other ships as well. In this scenario, there are two ships already berthed at the pier. Another DDG-51 is docked at the end of the pier while an FFG-7 is docked towards the land. These ships remain stationary throughout the scenario. Participants are required to moor their ship bow-in between these two ships. The environmental settings for this scenario were 0.8 knot on-setting wind and 0.3 knot on-setting current. One tug boat was available at the port for shiphandlers to use for maneuvering the ship during terminal stages. The tug always would attach to the bow on the ship's starboard side. The rationale for selecting this exercise was that ADOC students, who are our novice participants, were familiar with this exercise and for experts, this exercise requires skill and involvement. It also was chosen because it was simple enough to allow the novices a good chance to complete it successfully but was challenging enough to produce significant differences between the groups.

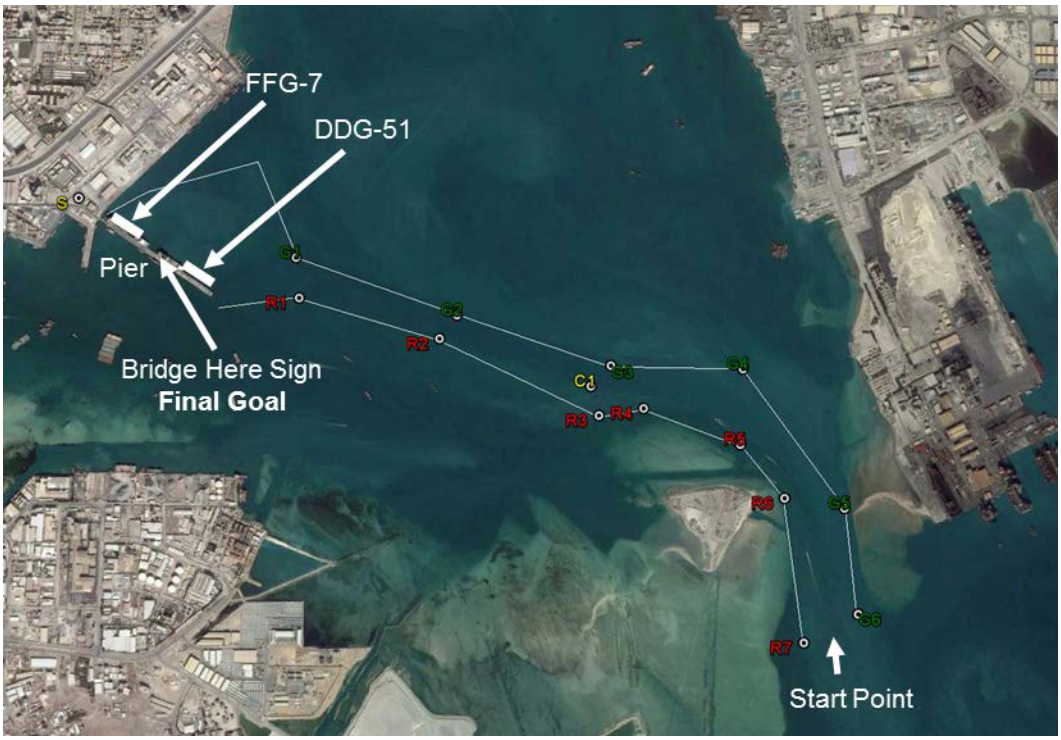


Figure 4. Approach Channel and Pier Setting of Port Mina Salman, Bahrain for the Simulation Exercise in COVE-3

The goals of this exercise for shiphandlers were to:

- Successfully navigate through the channel within the buoys
- Maintain safe speeds at all times as per requirement of place
- Maintain safe distance from a DDG-51 and FFG-7 at the pier
- Moor own ship in-between the other ships bow-in, maintaining safety parameters, and properly aligned with “Bridge Here” sign

E. RECRUITMENT PROCEDURES

The recruitment process involved my co-advisor sending bulk email from SWOS email exchange system to all the students and faculty at SWOS. All potential participants were requested to contact him for queries or scheduling in the experiments. My co-advisor went prior to the commencement of data collection at SWOS and facilitated the recruitment process. The participants were asked to give 60 minutes of their time for the experiment. The interested shiphandlers were then scheduled for the experiments.

F. SURVEYS

The participants were requested to fill out two surveys as part of the study. The survey forms are attached as Appendices A and B. A demographics survey provided information about participants' relevant shiphandling experience.

A post-task survey asked participants to gauge their performance and to list the strategies and visual cues used during different stages of the exercise. The stages mentioned for post-task survey were:

1. Approaching the berth
2. Getting the ship lined up at the berthing location
3. Bringing the ship close to the pier

G. PROCEDURES

The study started with setting up the scenario in the COVE system, configuring the ITS recording and setting up the eye tracking glasses. The participant would enter the COVE and the research team would brief the information on the consent form, stressing that participation was voluntary in nature and the participant could opt out of the experiment at any stage. The study team also discussed the equipment being used in the study and the type of data being recorded, including video and audio data. After providing written consent, the participant completed the demographics survey and then experimenters briefed the exercise using the chart table. The participant was then given time to study the chart, take notes and ask any queries about the setup. This would take approximately 15–20 minutes.

When the participant was ready to start the simulation exercise, experimenters fitted the eye tracking glasses with a comfortable nose piece and then conducted the eye tracking calibration. The calibration card was pinned up on a wall at an average height of a person and the participant was asked to focus his eyes onto the card's center point. The ETG controller software then performed the automatic calibration. Generally, the calibration procedure did not take more than a minute.

After calibration, the experimenters positioned the participant at the designated location in COVE. The participant was asked to view some specific object in the scene to validate the correct calibration. In most cases, we faced problems during this validation as the calibrated eye circle would show the participant's focus lower than where he reported actually looking. These observations were recorded for later processing during data extraction.

The simulation exercise would start when the participant gave a ready signal. He would then maneuver the ship through the narrow channel to moor at the pier. The time taken to complete the exercise varied between experts and novices. The experts would generally finish the scenario within 30 minutes, while some novices took more than an hour to finish the exercise. The participants were then asked to fill out a post-task survey and were thanked for taking their time to voluntarily participate in the study.

H. EXPERIMENT PHASE CONCLUSION

Upon completion of data collection, we inventoried and packed the eye tracker items for the trip back to NPS. The simulation logs, voice logs and ITS feedback were extracted from the COVE-ITS system and saved onto DVDs. All surveys, memory devices, and tablets were securely packed and were opened again upon reaching the designated project lab at NPS.

III. DATA PREPARATION

There are three principal data sources in our study beside surveys: eye-tracking data recorded on memory device in the recorder unit of ETG, Extensible Markup Language (XML) files containing simulator data, and voice logs containing speech commands given by the participants. Eye-tracking data was processed in the Tobii Pro Lab software through manual mapping of fixation information on an AOI chart. The data thus generated was analyzed for fixation metrics and visualization of attention patterns. The simulation and voice logs were converted from XML to comma separated value (CSV) format to be analyzed using the Pandas package of Python Language. The data type, extraction methods and tools used are described in this chapter.

A. EYE-TRACKING DATA

The data from Tobii Glasses 2 were recorded in proprietary format on the SD card. To extract the gaze data, we used the Tobii Pro Lab (Analyzer Edition) software running on the Windows 8 operating system. This software enables organization of all recordings in one project. Figure 5 shows the overview of the project created for extracting the eye-tracking data from the ETGs.

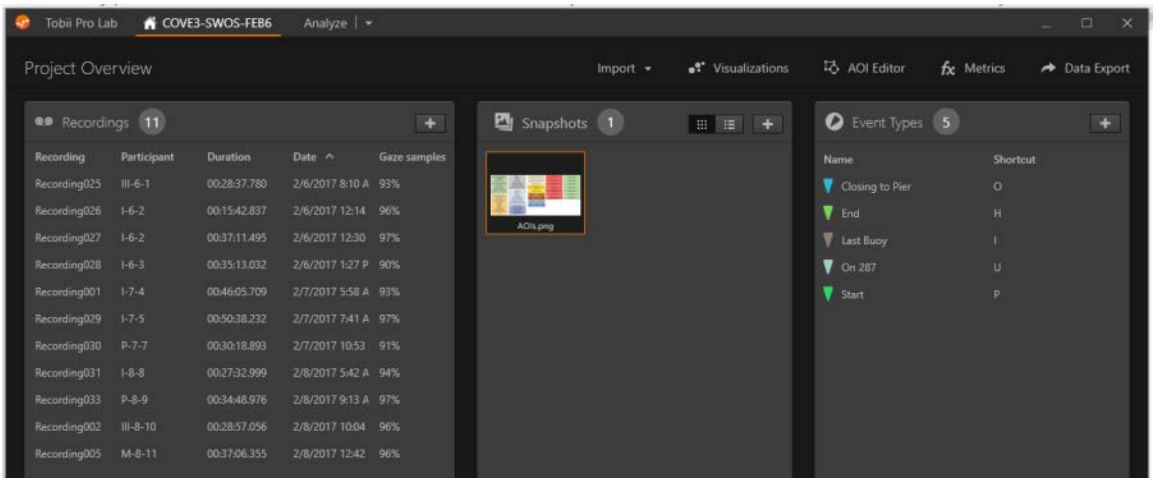


Figure 5. Tobii Pro Lab Software — Project Overview for Our Study's Data

On opening a specific participant's file, the system displays the recorded video in the application, as shown in Figure 6. The circle superimposed on the video shows the gaze position of the participant. The controls for play/pause/play in compressed time and play frame-by-frame are given in the bottom left of the video. Unlike video players, the frames in Tobii Pro Lab are distinguished based upon the type of gaze filter in use.



Figure 6. Recorded Video Display in Tobii Pro Lab with Superimposed Gaze Position

There are two preset filters in Tobii Pro Lab to view the gaze data. It can be set in the "Gaze Data" pane available on the right side of the application's graphical user interface. In the "Gaze Filter" selection, "Raw" means no filter is selected and the circle is updated at 50 Hz intervals for each eye. The other two selections are Tobii I-VT (Fixation) and Tobii I-VT (Attention) filters. VT means use of velocity threshold method for classification of eye movement. For this study, we used Tobii I-VT (Attention) filter with the settings shown in Table 3.

Table 3. Parameters of Tobii I-VT (Attention) Filter (Courtesy or Tobii AB). Adapted from Tobii Pro Lab software.

Gaze filter	Tobii I-VT (Attention)
Gap fill-in (interpolation)	Off
Noise Reduction	Moving median
Window size (samples)	3
Velocity calculator - Window length(ms)	20
I-VT classifier - Threshold(°/s)	100
Merge adjacent fixations	On
Max time between fixations (ms)	75
Max angle between fixations (°)	0.5
Discard short fixations	On
Minimum fixation duration (ms)	60 ms

Eye tracking experiments use AOIs to describe where a participant has fixated his eyes in a particular frame. Generally, in eye tracking studies the subject's viewpoints remain static throughout the experiment. Therefore, analysts can define AOIs on an image of the recorded environment and use the automatic mapping feature of Tobii Pro Lab to extract information regarding fixation time and areas. In our study, however, because the ship's movement caused the participant's view to continuously change, the AOIs were also changing with time. Hence, we could not use the automatic mapping feature. To overcome this problem, we devised a method to manually extract the AOIs from the videos. We subjectively analyzed the experts and novices' videos to ascertain different areas where they fixated their gaze and then designed an initial AOI chart. We improved this AOI chart in iterations after testing it over different segments of videos for its usability. We imported the final chart as shown in Figure 7 into the AOI editor of

Tobii Pro Lab. In sum, 49 AOIs were identified. The description of these AOIs is in Table 4.

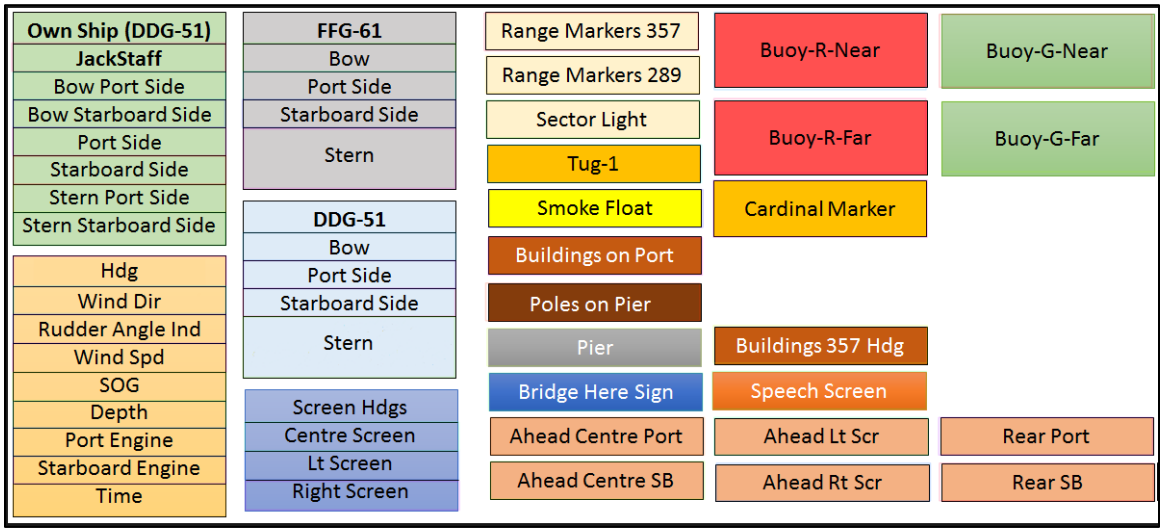


Figure 7. Areas of Interest Defined to Map Fixation Data

Table 4. Description of AOIs Used in Eye Tracking

AOI	Description
Ahead_Cen_P (Ahead Center Port)	Looking within left half of center screen but not at any specific object
Ahead_Cen_SB (Ahead Center Starboard)	Looking within right half of center screen but not at any specific object
Ahead_Scr_Lt (Ahead Lt Scr)	Looking at left screen but not at any specific object
Ahead_Scr_Rt (Ahead Lt Scr)	Looking at right screen but not at any specific object
Bridge_Here_Sign	"Bridge Here" sign on the pier
Building_357Hdg	Buildings visible on start leg where start heading is 357°
Buildings (Building on Port)	Buildings on shore at the pier side
Buoy_G_ Far	Green buoys ahead of ship's current position (not to include the closest)
Buoy_G_Near	Green buoy closest to current position of ship
Buoy_R_ Far	Red buoys ahead of ship's current position (not to include the closest)

AOI	Description
Buoy_R_Near	Red buoy closest to current position of ship
Cardinal_Marker	Marker with yellow light in center of channel
DDG_Bow (DDG-51 Bow)	Bow of DDG-51 berthed at pier
DDG_P (DDG-51 Port Side)	Port side of DDG-51 berthed at pier
DDG_S (DDG-51 Starboard Side)	Starboard side of DDG-51 berthed at pier
DDG_Stern (DDG-51 Stern)	Stern of DDG-51 berthed at pier
FFG_Bow (FFG-61 Bow)	Bow of FFG-61 berthed at pier
FFG_P (FFG-61 Port Side)	Port side of FFG-61 berthed at pier
FFG_S (FFG-61 Starboard Side)	Starboard side of FFG-61 berthed at pier
FFG_Stern (FFG-61 Stern)	Stern of FFG-61 berthed at pier
Inst_Depth (Depth)	Fathometer
Inst_Engine_P (Port Engine)	Instrument panel port engine indicator
Inst_Engine_S (Starboard Engine)	Instrument panel starboard engine indicator
Inst_Hdg (Hdg)	Instrument panel current heading indicator
Inst_Rudd_Ang (Rudder Ang Ind)	Instrument panel rudder angle indicator
Inst_SOG (SOG)	Instrument panel speed over ground indicator
Inst_Time (Time)	Instrument panel time
Inst_Wind_Dir (Wind Dir)	Instrument panel wind direction indicator
Inst_Wind_Spd (Wind Spd)	Instrument panel wind speed indicator
Own_Bow_P (Ownship-Bow Port Side)	Left half of own ship bow
Own_Bow_S (Ownship-Bow Starboard Side)	Right half of own ship bow
Own_Jackstaff (JackStaff)	Own ship jackstaff
Own_P	Own ship port side

AOI	Description
(Port Side)	
Own_S (Starboard Sie)	Own ship starboard side
Own_St_P (Stern Port Side)	Own ship stern (left side)
Own_St_S (Stern Starboard Side)	Own ship stern (right side)
Pier	Pier
Rear_P (Rear Port)	Rear projection screen left half
Rear_SB (Rear SB)	Rear projection screen right half
Rng_357 (Range Markers 357)	Range markers at 357° heading
Rng_289 (Range Markers 289)	Range markers at 289° heading
Scr_Speech (Speech Screen)	Small screen where speech commands appear
ScrHdg_Cen (Screen Hdg Center Screen)	Heading at center screen
ScrHdg_Lt (Screen Hdg Lt Screen)	Heading at left screen
ScrHdg_Rt (Screen Hdg Right Screen)	Heading at right screen
Sector_Light	Sector light
Smoke_Float	Smoke float
Things_Pier (Poles on Pier)	Objects at pier
Tug_1	Tug boat

Note: Some AOI cells have two names. The top names are the ones used in the graphs in Chapter IV, while the names in parentheses indicate those used in Figure 7 of this thesis.

We manually mapped the fixation position on the AOI chart for each frame in participants' recordings. During this process, ETG's lower gaze reporting errors during the calibration process were adjusted by choosing the AOI above the actual indicated gaze point. Figure 8 shows a screen shot of manual mapping of gaze circle to AOI chart to the right.

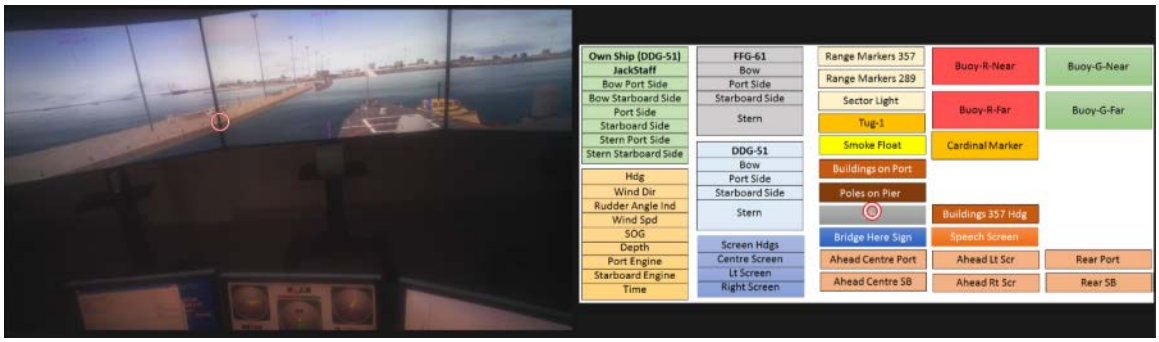


Figure 8. Marking the Fixation Information from Video on the AOIs Chart

The mooring exercise has different phases and each phase requires application of different strategies and visual cues. Hence, each participant's recording was segmented for extraction of data. After manually mapping the gaze data on relevant AOIs, five events were defined as shown in Figure 9.

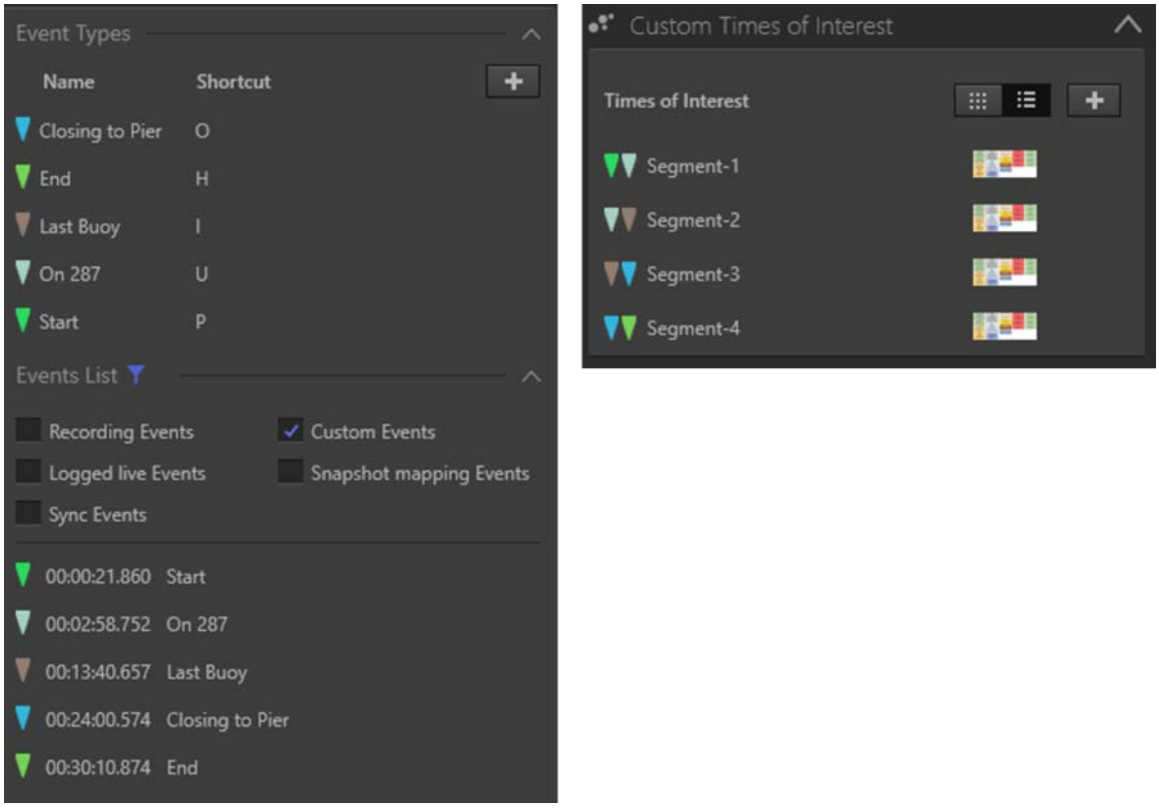


Figure 9. Events Definition on Top and as Marked in a Recording at the Bottom

These events are defined as follows:

- **Start.** Start of exercise, each participant will maintain heading 357° until turning left into the channel.
- **On 287.** The participant has made first left turn and is steady in the channel. Track on this leg is 287°.
- **Last Buoy.** The participant is crossing the last pair of buoys in the channel.
- **Closing to Pier.** The participant ship's stern has cleared the berthed DDG-51 and now has to close-in to the pier.
- **End.** End of exercise.

The times of interest were defined as the segments between consecutive events. Four such segments were thus defined as shown on the right in Figure 9. For each segment, eye gaze metrics and raw data were exported for each participant's recording in the study. This data was then converted into a CSV file for further analyses. The description of extracted eye gaze data is as follows:

1. Eye-Gaze Metrics

The Tobii Pro Lab can export a number of metrics after automatic/manual mapping of data onto snapshots. In our observational study, since we used attention filter for processing the gaze data, the relevant metrics available from the software are listed as follows:

- **Interval Duration.** The time spent by each participant in each of four event-based time segments defined above
- **AOI Fixation Count.** The number of fixations within each AOI
- **AOI Total Fixation Duration.** The total time each participant has fixated in each AOI
- **AOI Average Fixation Duration.** The participants' average duration of fixation in each AOI

Tables 5 and 6 illustrate some of this output.

Table 5. Total Fixation Duration (Seconds) for Some AOIs in Segment 4 (Sample)

Participant	Ahead_Cen_P	Ahead_Cen_SB	Ahead_Scr_Lt	Ahead_Scr_Rt	Bridge_Here_Sign	Buildings
I-7-4	16.98	7.10		1.92	10.86	3.66
III-8-10	23.22	0.54			35.21	0.42
M-8-11	3.42		0.10			
III-6-1	28.52	0.78			6.00	0.26
I-6-2	287.07	33.93	2.58	0.56	30.05	10.14
I-6-3	41.54	5.84	1.40	1.00	15.66	12.08
I-7-5	110.52	68.53	117.10	2.24	2.68	10.12
P-7-7	2.96	0.38		0.58	48.53	
I-8-8	1.94	1.96			27.65	1.00
P-8-9	34.49	6.10	0.16	1.04	38.34	16.36

Table 6. Total Fixation Count for Some AOIs in Segment 4 (Sample)

Participant	Ahead_Cen_P	Ahead_Cen_SB	Ahead_Scr_Lt	Ahead_Scr_Rt	Bridge_Here_Sign	Buildings
I-7-4	39	19		8	30	8
III-8-10	48	3			45	1
M-8-11	5		1			
III-6-1	36	3			10	2
I-6-2	366	35	7	1	33	23
I-6-3	177	31	10	7	63	61
I-7-5	181	72	145	15	2	17
P-7-7	7	1		1	80	
I-8-8	11	6			76	3
P-8-9	42	12	1	6	55	21

Because the time taken to complete each segment varies between participants and is quite large for novices, both fixation count and duration metrics were converted to respective percentage of duration/count within each AOI for all participants. This was done by importing the CSV metrics file into Python and using the Pandas package to do the conversion.

2. Raw Data Export

The other type of data available from the Tobii Pro Lab is the raw data recorded at 50 Hz. A number of parameters can be exported by the software. For our observational

study, however, we have mapped the fixation information on an AOI chart. The following parameters were of our interest in analysis (see Table 7):

- **Recording Timestamp.** Timestamp counted from the start of the recording (t0 = 0) in milliseconds
- **Pupil Diameter Left.** Estimated size of the left eye pupil in millimeters
- **Pupil Diameter Right.** Estimated size of the right eye pupil in millimeters
- **Eye Movement Type.** Type of eye movement classified by the fixation filter. The possible values are (1) Fixation, (2) Saccade, (3) Eye-NotFoundMovement, (4) UnknownEyeMovement.
- **Gaze Event Duration.** The duration of the currently active eye in milliseconds
- **AOI hit [Snapshot Name – AOI Name].** Reveals if there is a fixation within a given AOI on a given Snapshot. The value can be 0 or 1.

Table 7. Example of Exported Raw Data Information

Recording	Pupil diameter left	Pupil diameter right	Eye movement	Gaze event	Eye movement	AOI hit [AOIs.p	AOI hit [A	AOI hit [A	AOI hit [A	AOI hit [A
123882	5.144023418	5.422325611	Fixation	800	200	0	0	0	0	0
123902	5.141446114	5.42331028	Fixation	800	200	0	0	0	0	0
123922	5.141457558	5.423253536	Fixation	800	200	0	0	0	0	0
123942	5.151737213	5.426864147	Fixation	800	200	0	0	0	0	0
123962	5.180847168	5.405575275	Fixation	800	200	0	0	0	0	0
123982	5.20223856	5.394486904	Fixation	800	200	0	0	0	0	0
124002	5.227672577	5.400628567	Fixation	800	200	0	0	0	0	0
124022	5.246839523	5.400279522	Fixation	800	200	0	0	0	0	0
124042	5.261465073	5.388763428	Fixation	800	200	0	0	0	0	0
124062	5.274790764	5.377164841	Fixation	800	200	0	0	0	0	0
124082	5.300327778	5.366070271	Fixation	800	200	0	0	0	0	0
124102	5.32062912	5.349966049	Fixation	800	200	0	0	0	0	0
124122	5.337347984	5.33462286	Fixation	800	200	0	0	0	0	0
124142			EyesNotFound	140	87	0	0	0	0	0
124162			EyesNotFound	140	87	0	0	0	0	0
124182			EyesNotFound	140	87	0	0	0	0	0
124202			EyesNotFound	140	87	0	0	0	0	0
124222			EyesNotFound	140	87	0	0	0	0	0
124242			EyesNotFound	140	87	0	0	0	0	0
124262			EyesNotFound	140	87	0	0	0	0	0
124282	5.372200489	5.367965698	Unclassified	20	27	0	0	0	0	0

The file format of Tobii Pro Lab raw data export is tab separated values ‘tsv’. This data was imported into Python for further processing using data analysis and visualization packages. There were four such files for each participant (one file for each time segment) and only fixation information was extracted from raw data for further analyses.

B. SIMULATOR LOGS

The simulator logs are produced by the COVE system and extracted via the ITS. These files are named as “covestate_time_date.log” and are in XML format. The information contained in these files can broadly be categorized as:

- Timestamp
- Objects
- Attributes
- Values

The timestamp is in epoch time format from launch of the scenario until the end of the recording. There are more than seventy objects in a simulator log. A log entry is added at one-second intervals for an object if its parameters change in the scenario. Each object may have many attributes; the ship object has position, velocity, acceleration, depth of water, etc., as just a few of its attributes. Each attribute has its value stored at that timestamp. Figure 10 shows one of these logs.

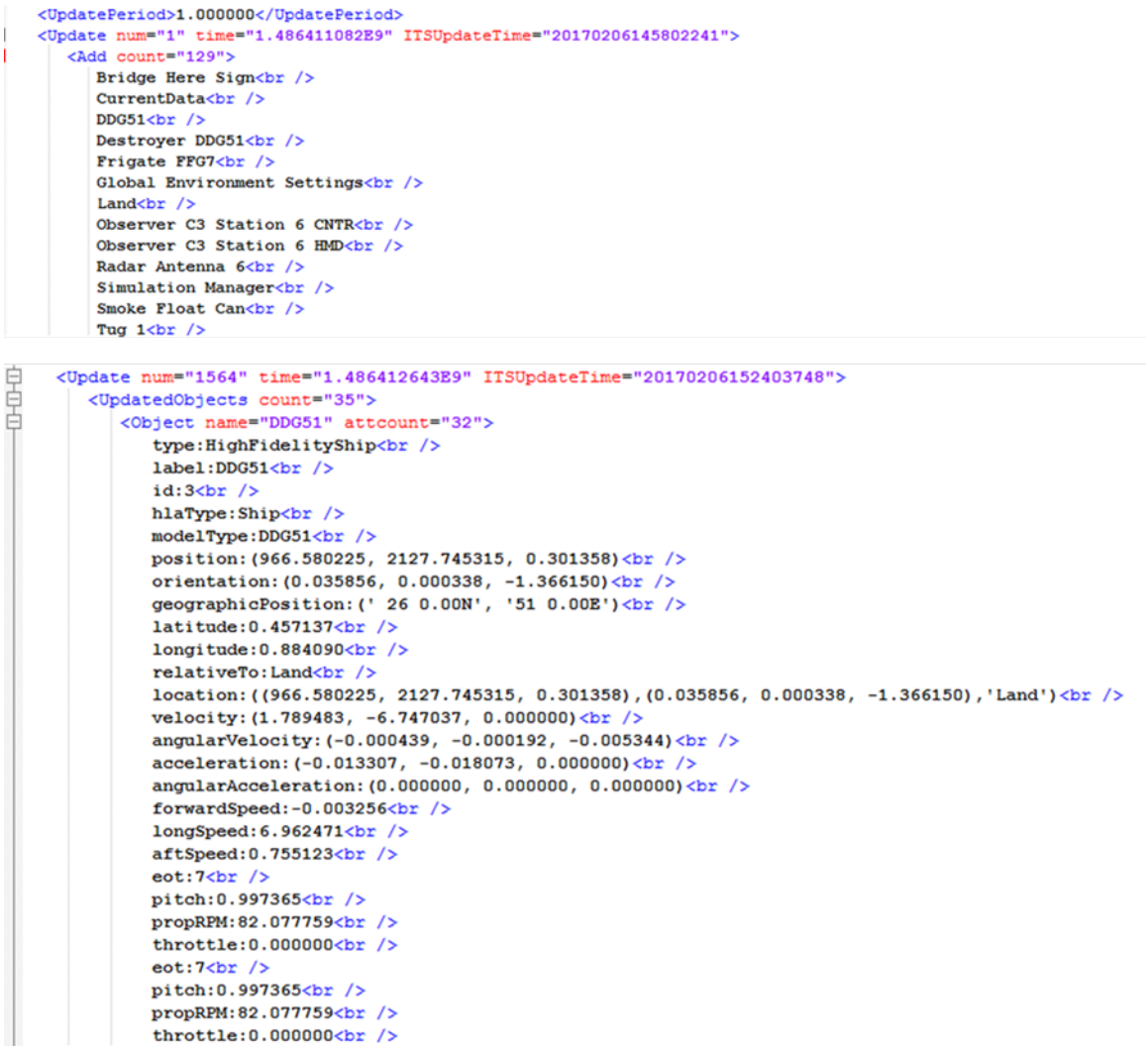


Figure 10. Simulator Log File Objects and Attributes

The most important object for our analysis is the own-ship (i.e., “DDG51”) object in the simulator log. The information from these logs was extracted in Python using the “etree” and “re” packages, which interface with XML files and regular expressions, respectively. The data was then incorporated into the Pandas dataframe, where each row corresponds to DDG51’s parameters at a certain timestamp, and was also exported into CSV files. Table 8 shows data in one of these files.

Table 8. Extracted Data Example From Simulator Log Files in CSV Format

	orientation	BrgBridge	distBridge	distDestrc	distFrigat	longSpee	forwardSp	aftSpeed	SteeringA	pitch1	throttle1	pitch2	throttle
1340	283.9232369	1.623373	795.9479	579.2087	953.6002	14.29758	-0.22243	-0.10753	0	0.960217	0.160111	0.960217	0.1601
1341	283.8997457	1.62379	789.698	572.9873	947.4394	14.20464	-0.22243	-0.10753	0	0.949611	0.160111	0.949611	0.1601
1342	283.8799786	1.624644	777.1986	560.547	935.1218	14.08031	-0.22243	-0.12761	0	0.939006	0.160111	0.939006	0.1601
1343	283.862618	1.616043	770.6017	555.1612	927.7331	13.9893	-0.22243	-0.14823	0	0.931494	0.160111	0.931494	0.1601
1344	283.8477211	1.616413	764.3495	548.9527	921.5694	13.86519	-0.22243	-0.14823	0	0.920888	0.160111	0.920888	0.1601
1345	283.8351733	1.617174	751.8455	536.5392	909.2459	13.68227	-0.22243	-0.16856	0	0.869579	0.160111	0.869579	0.1601
1346	283.8246882	1.608214	745.2998	531.2966	901.873	13.55557	-0.22243	-0.16856	0	0.830102	0.160111	0.830102	0.1601
1348	283.816323	1.608532	739.0456	525.1043	895.7063	13.40036	-0.22243	-0.18827	0	0.779658	0.160111	0.779658	0.1601
1349	283.8097913	1.608854	732.7915	518.9136	889.5409	13.19861	-0.22243	-0.18827	0	0.727021	0.160111	0.727021	0.1601
1350	283.8053795	1.609182	726.5374	512.7246	883.3767	13.02949	-0.22243	-0.20824	0	0.680964	0.160111	0.680964	0.1601
1351	283.8033169	1.609856	714.0295	500.3517	871.0526	12.84754	-0.22243	-0.20824	0	0.632714	0.160111	0.632714	0.1601
1352	283.8036606	1.600351	707.5349	495.2989	863.682	12.64547	-0.22243	-0.22849	0	0.580077	0.160111	0.580077	0.1601
1353	283.8064681	1.600615	701.279	489.1314	857.5148	12.44946	-0.22243	-0.24821	0	0.529633	0.160111	0.529633	0.1601

C. VOICE LOGS

The other data of interest were voice commands spoken by the shiphandler. Conning officers on a Navy ship don't actually control the rudder or engines, instead giving orders via voice to a helmsman and lee helmsman. Similarly, voice commands are the only way for shiphandlers to interact with the COVE system, which are interpreted through COVE's voice recognition system. All this data is recorded with timestamps in customized XML format. An example is shown in Figure 11.

```
TIME=20170207134042859|CHAN=1|EVNT=SWIclst|VALU=Session 20170207134042-C3VOICE6-001-LOG started|SRC=SWIrec|UCPU=0|SCPU=0
TIME=20170207134050187|CHAN=1|EVNT=EXERCISE_STATUS|VALUE=Paused
TIME=20170207134134078|CHAN=1|EVNT=EXERCISE_STATUS|VALUE=Running
TIME=20170207134136093|CHAN=1|EVNT=EXERCISE_STATUS|VALUE=Paused
TIME=20170207135359375|CHAN=1|EVNT=EXERCISE_STATUS|VALUE=Running
TIME=20170207135429281|CHAN=1|EVNT=cove_ptr_press
TIME=20170207135430343|CHAN=1|EVNT=cove_ptr_release
TIME=20170207135430890|CHAN=1|EVNT=SWIrslt|MEDIA=application/x-vnd.speechworks.emma+xml;strictconfidencelevel=1;mrcpv=2.
<?xml version='1.0'?><result>
<interpretation grammar="file://C:/Program%20Files/Nuance/SpeechFed/grammars/ddg51_wrapper.grxml?SWI_vars.current_headin
<input mode="speech">left_standard rudder</input>
<instance>
<command_1 confidence="35">simple_rudder_direction</command_1>
<SWI_meaning>left standard rudder</SWI_meaning>
<command_type_1 confidence="35">simple</command_type_1>
<command confidence="35">simple_rudder_direction</command>
<auto_course_1 confidence="35">not_set</auto_course_1>
<rudder_direction_1 confidence="22">left</rudder_direction_1>
<rudder_intensity_1 confidence="35">standard</rudder_intensity_1>
<approved_command_1 confidence="35">yes</approved_command_1>
<rejection_threshold_1 confidence="35">30</rejection_threshold_1>
<disallow_1 confidence="35">0</disallow_1>
<literal_1 confidence="22">left_standard rudder</literal_1>
<commands_parsed confidence="35">1</commands_parsed>
<SWI_literal>left_standard rudder</SWI_literal>
<SWI_grammarName>file://C:/Program%20Files/Nuance/SpeechFed/grammars/ddg51_wrapper.grxml?SWI_vars.current_heading=356&am
</instance>
</interpretation>
</result>
TIME=20170207135430890|CHAN=1|EVNT=SWIDSCN|RSLT=ACCEPT|WVNM=C:\Program Files\Nuance\SpeechFed\data\2017\02\07\2017020713
TIME=20170207135430890|CHAN=1|EVNT=PROMPT_START|ID=04A56658|TEXT=left standard rudder, aye sir.
```

Figure 11. Voice Log Files Format

Voice commands are extracted in Python using ‘etree’ and ‘re’ packages and converted into a table with timestamps as index. The timestamps are in seconds. Table 9 shows the format of these extracted voice commands.

Table 9. Extracted Table Format of Voice Commands

Timestamp	command	auto_course	engine_designation	engine_direction	engine_speed	percent_pitch	rudder_degrees	rudder_direction	rudder_int
2086	observer_position	not_set							
2090	rudder_amidships	not_set							
2092	rudder_amidships	not_set							
2159	simple_rudder_direction	not_set						right	full
2162	simple_rudder_direction	not_set						left	full
2166	tug_direction	not_set							
2177	simple_rudder_direction	not_set						left	hard
2181	tug_direction	not_set							
2197	engine_stop	not_set	all						
2199	tug_status	not_set							
2219	engine_command	not_set	port	back	one_third				
2227	tug_direction	not_set							
2249	rudder_amidships	not_set							
2265	engine_choice_and_pitch	not_set	port	back		20			
2309	tug_status	not_set							
2325	engine_command	not_set	port	back	one_third				
2334	engine_choice_and_pitch	not_set	port	back		22			
2383	tug_direction	not_set							
2397	tug_status	not_set							
2409	simple_rudder_direction	not_set						right	standard
2418	tug_direction	not_set							
2427	engine_stop	not_set	all						
2430	tug_status	not_set							
2434	rudder_amidships	not_set							
2461	engine_command	not_set	port	ahead	one_third				
2466	engine_choice_and_pitch	not_set	starboard	back		22			
2498	simple_rudder_direction	not_set						right	standard
2508	tug_direction	not_set							
2512	engine_command	not_set	starboard	back	one_third				
2522	rudder_amidships	not_set							
2525	tug_status	not_set							
2529	engine_choice_and_pitch	not_set	starboard	back		20			
2532	tug_direction	not_set							
2549	tug_status	not_set							

There are 14 attributes found in the voice logs. Depending upon the command, one or more attributes in the following list will have corresponding values in the table.

1. Auto course
2. Engine designation
3. Engine direction
4. Engine speed
5. Observer location
6. Percentage pitch
7. Rudder degrees
8. Rudder direction

9. Rudder intensity
10. Speed
11. Tug command
12. Tug direction
13. Tug number
14. Tug speed

D. DATA SYNCHRONIZATION

Simulator data, voice logs and eye-tracking data are timestamped with their own machine clock reference. Hence, they all had to be manually synchronized together. For eye-tracking data, the videos were monitored in Tobii Pro Lab to note the timing of a particular rudder command. Then in voice logs, the timestamp of that rudder command was noted from the extracted table of voice commands. Further, in simulation data the rudder column was monitored for change in rudder direction as per the voice command. These three time references were noted for all the participants' recordings and then synchronization was incorporated in all the extracted data files for data analyses.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. ANALYSES AND RESULTS

In order to address the research questions, several steps were completed. We first tried to establish expertise differences in shiphandling exercise performance by positional and time references. Then, eye tracking metrics were analyzed to test whether attention allocation differences occur between the experts and novices. Additionally, scan transition graphs and timeline graphs for AOIs were built to explore expertise differences in strategies. We also tried to evaluate novices' performance using experts' data in conjunction with analytical hierarchy process (AHP). Analyses and results are presented in the order outlined above; results from the post-task survey also are illustrated. The analytical tools used included box plots of eye tracking metrics to compare experts' and novices' fixation in all AOIs. Further, important differences between groups were explored using both parametric (two sample *t*-tests) and non-parametric tests (Mann-Whitney U test). Network graphs and timelines were used to explore the scanning techniques of individual participants.

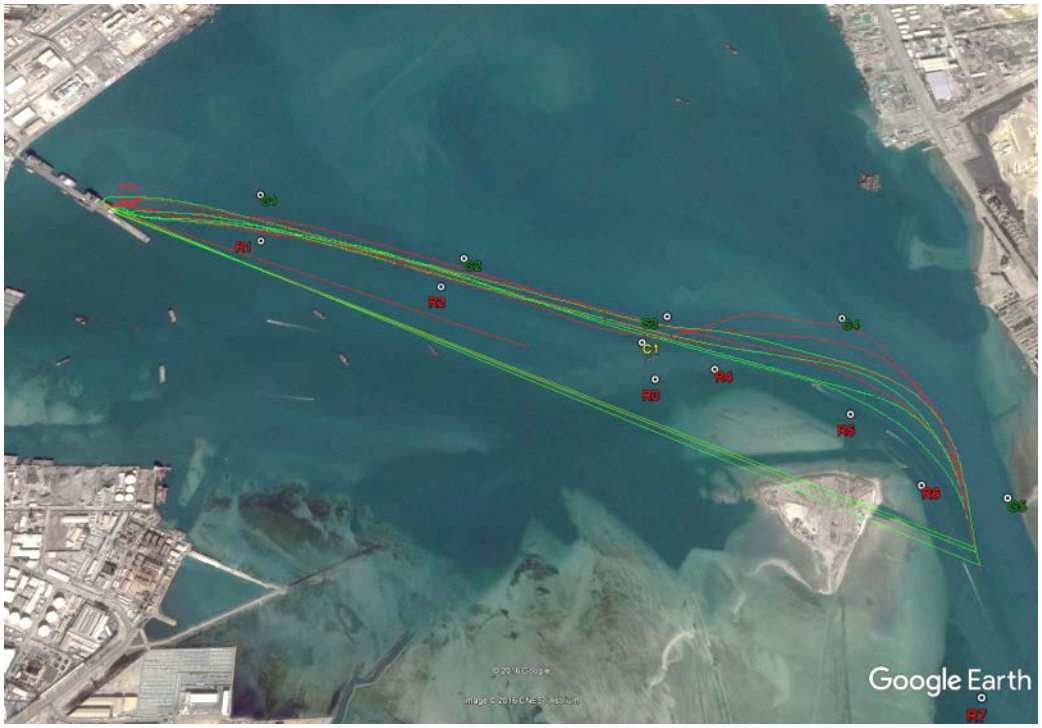
A. EXPERTISE DIFFERENCES IN SHIPHANDLING PERFORMANCE

Prior to examining whether certain Attention-allocation patterns are associated with better shiphandling performance, we first confirmed that the experts did in fact perform better on the simulation exercise. We examined both the ship's path and time to complete the exercise. For analysis purposes, the simulation exercise is divided into four segments as per sponsor's input. A brief description of this division is as following:

- **Segment 1**: From start point to finishing the first turn on heading 287°
- **Segment 2**: Steady course inside the channel until reaching last pair of buoys
- **Segment 3**: From last pair of buoys until stern has cleared the berthed DDG-51
- **Segment 4**: Clearing the berthed DDG-51 until moored

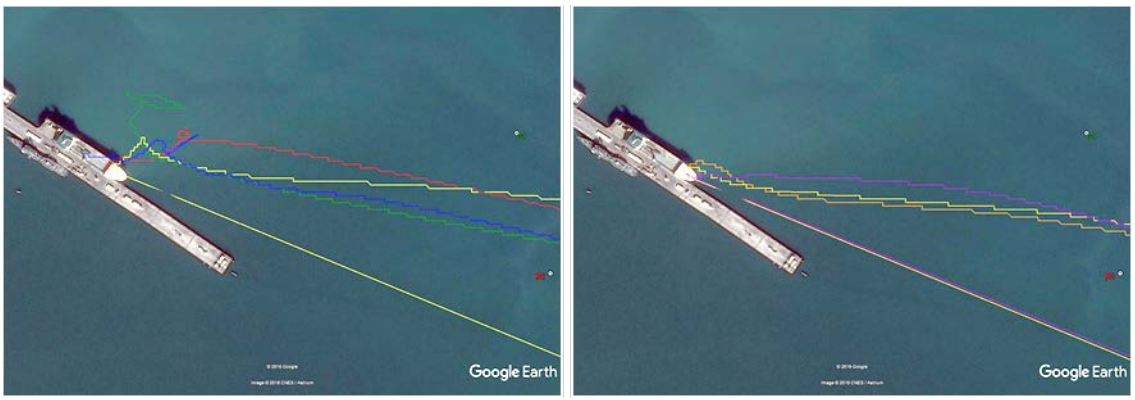
1. Ship's Path Visualization

The ship's positional information is available in geographic coordinates (latitude/ longitude) format in simulation logs. This information was extracted and processed in Keyhole Markup Language to visualize in Google Earth as shown in Figure 12. The green color indicates experts' paths and red color is used for novices' paths. The difference between the two groups is visible after the first turn into the channel where two red paths have overshot the middle of channel. By the cardinal marker (C1), however, every path is almost aligned taking reference from the sector light, and segment 2 positional reference appears the same for both groups. During segments 3 and 4, the difference of paths becomes more evident as shown in Figure 13. It also illustrates the increased variability in novices' paths.



Path traced out by experts (green) and novices (red) during simulation exercise

Figure 12. Ships' Path in Simulation Exercise



Novice (left) and expert (right) paths during segments 3 and 4 highlighted in colors

Figure 13. Participants' Paths during Segments 3 and 4

2. Time Taken by Participants

The time taken by the participants to complete the simulation exercise ranges between 26 and 49 minutes. As expected, time to complete was longer and more variable among the novices than experts for overall time and time to complete each segment (see Figure 14).

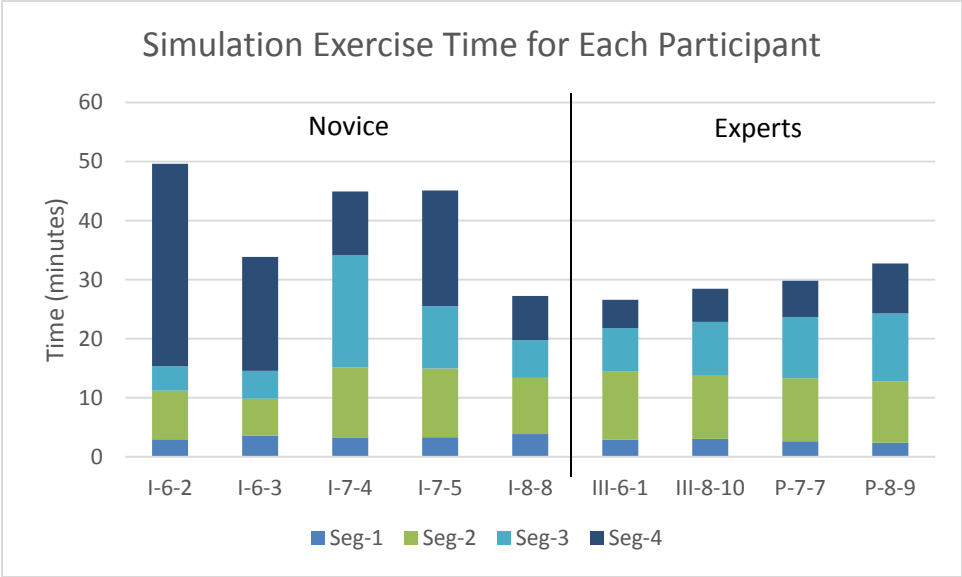


Figure 14. Time Taken by Each Participant to Complete the Exercise (Segmented)

As shown in Figure 15, experts showed almost no variability in segments 1 and 2, while in segments 3 and 4 there is little variability. The median time in segments 1 and 4 is less for experts than novices, while it is slightly more in segment 2 and more pronounced in segment 3. This result indicates that novices maintained more speed in segment 3 than the experts (from last buoys until preparing for landing after clearing own stern from berthed DDG-51). In segment 4, which involves controlling the ship's lateral movement to moor the ship, novices struggled and took a lot more time. Thus, the time and paths information confirm that expertise differences in shiphandling performance occurred.

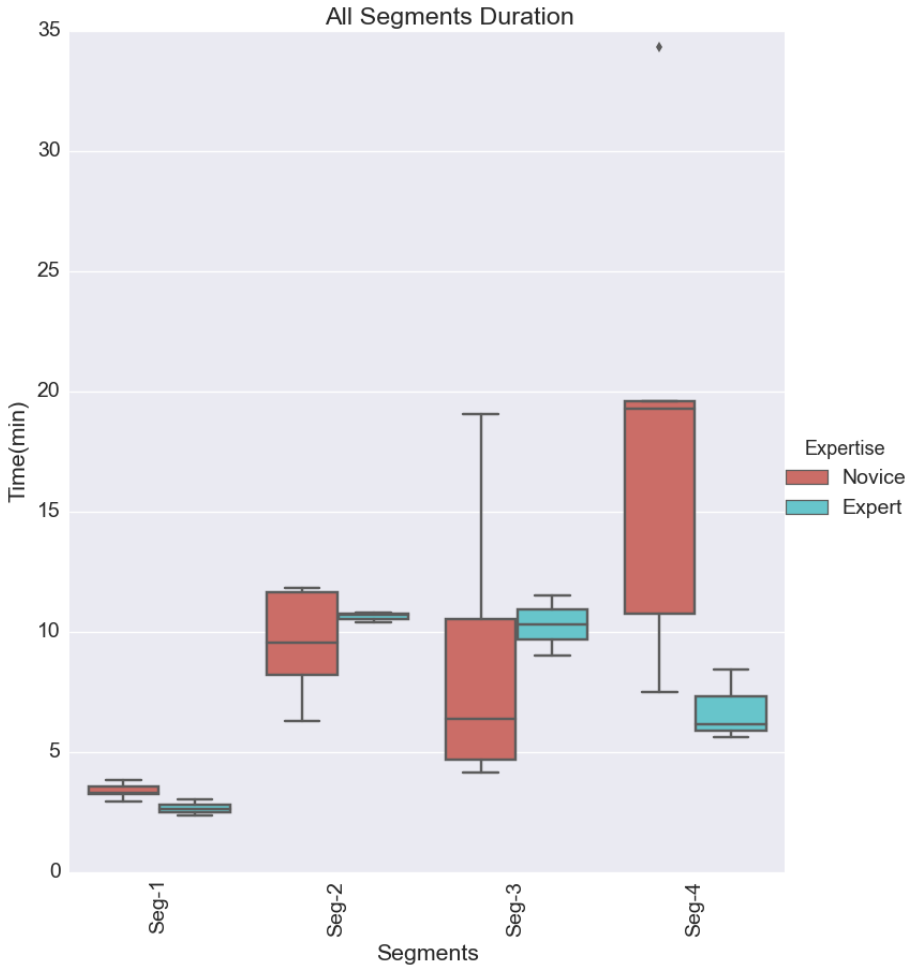


Figure 15. Time Taken to Complete the Exercise (Segmented and Grouped)

B. EYE-TRACKING FIXATION METRICS

Eye tracking analyses were broken into two parts. The first part examined general attention allocation metrics, such as the percent of time participants spent looking at each AOI for segment 3 and segment 4. The second part examined scan transitions, which include network graphs having nodes as AOIs and edges indicating fixation transitions between AOIs. The network graph is supported by AOI timeline graphs that show the participant's fixation on AOIs distributed over time.

1. General Attention Allocation Metrics

The eye tracking metrics were extracted after processing data in Tobii Pro Lab. These metrics are AOI fixation count, AOI total fixation duration and AOI average fixation durations in each segment. Because each participant took a different time to complete the segments, the absolute counts and total fixation durations have been converted into percentage count/duration for each AOI. Further, the data was aggregated for experts and novices.

The sponsor has conveyed interest in analyzing segment 3 and segment 4, which are the most challenging and error prone; therefore, the following discussion presents analysis of these segments only. For segment 1 and segment 2, the metrics are given in Appendix C.

Single-sided independent sample t-tests with unequal variance and Mann-Whitney U-Test were performed on the two groups across all AOIs. Because this is a pilot study, an alpha level of .10 is implemented. Results with p-values < 0.1 for MW U-Test are reported in the following paragraphs along with independent sample t-test for reference.

a. Segment 3 Metrics

(1) AOI Fixation Count Percentage

In segment 3, expertise differences in the percent of fixation counts spent were seen in three AOIs: the left half of the center screen where the pier is now visible; the jackstaff, and the left half of the stern view. Experts predominantly focus on the left half

of the center screen where the pier is now within good visual range and they are planning for the landing while looking at the conditions at the pier as shown, in Figure 16.

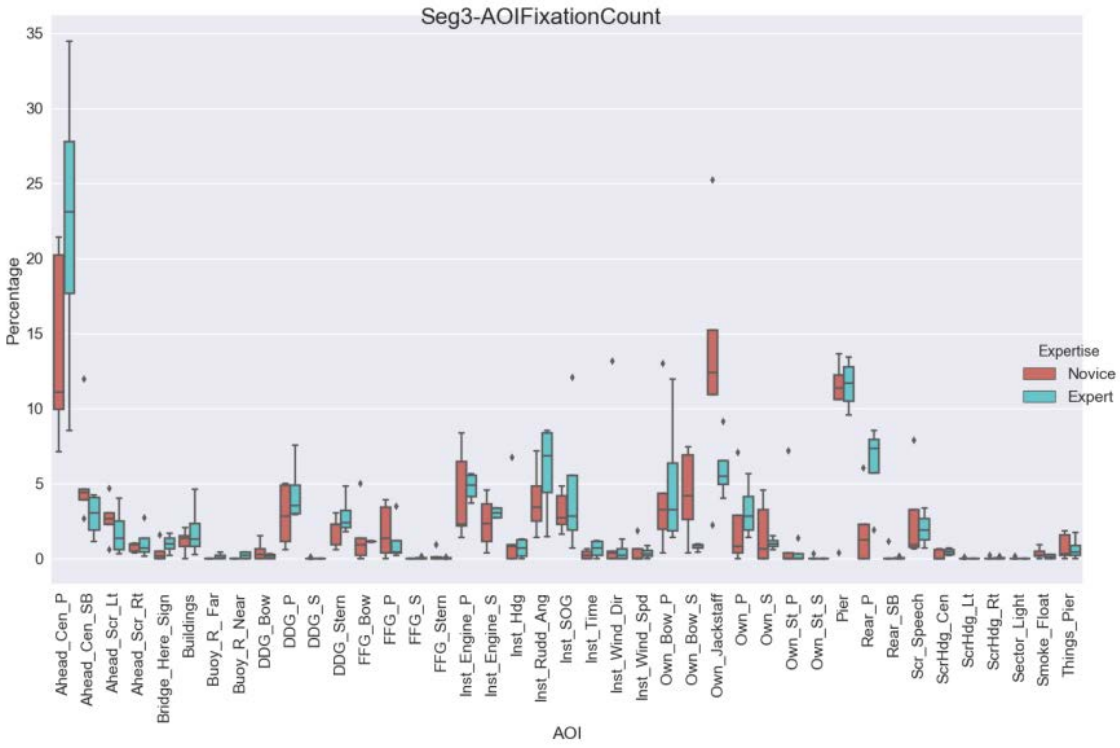


Figure 16. AOI Fixation Count Percentages in Segment 3

Table 10 show that experts have fixated more on the left half of the stern view screen than novices. In contrast, the AOIs where experts' fixation count percentage is less than the novices are the starboard side of own-ship bow and the jackstaff.

Table 10. T-test and MW U-test Result for Seg-3 Fixation Count Percentage

AOI	T-Statistic	p-value (T-test)	MW U-Statistic	p-value (MW Test)
Rear_P	2.335	0.029	18	0.032
Bridge_Here_Sign	1.265	0.123	16	0.089
Ahead_Cen_SB	-1.440	0.102	4	0.089
Own_Bow_S	-2.639	0.028	4	0.089
Own_Jackstaff	-1.853	0.063	4	0.089

(2) AOI Total Fixation Duration Percentage

In Figure 17, the difference between groups is even more obvious for total fixation duration percentage on the left half of the center screen. Experts have spent more time fixating here while novices have fixated more on the jackstaff. Experts also fixated longer on the Instrument Panel Port Engine and Starboard Engine than the novices.

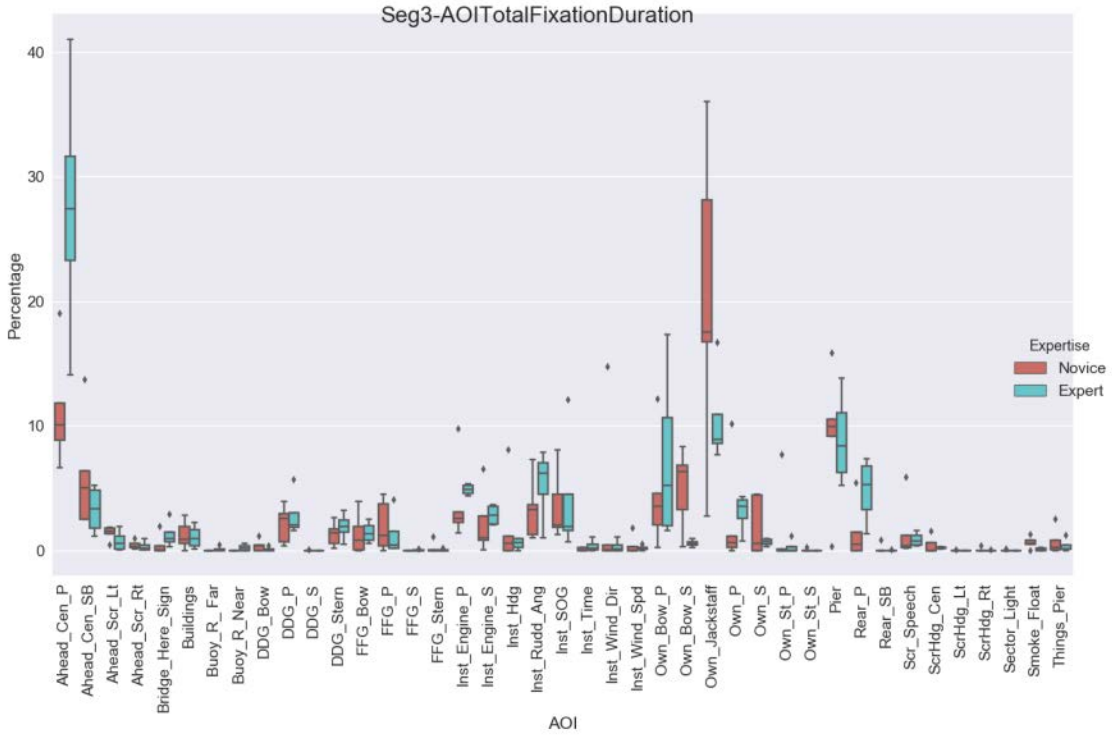


Figure 17. AOI Fixation Duration Percentages in Segment 3

Table 11 shows that experts have more total fixation duration percentage on the left half of the center screen and the left half of the stern view. The AOIs where experts have less total fixation duration percentage are own-ship jackstaff, the starboard side of the bow and the smoke float for wind estimation.

Table 11. T-test and MW U-test Result for Seg-3 Total Fixation Duration Percentage

AOI	T-Statistic	p-value (T-test)	MW U-Statistic	p-value (MW Test)
Ahead_Cen_P	2.742	0.027	19	0.019
Rear_P	1.949	0.049	17	0.055
Bridge_Here_Sign	1.198	0.140	16	0.088
Inst_Engine_P	0.714	0.256	16	0.089
Own_Jackstaff	-1.611	0.084	4	0.089
Own_Bow_S	-3.063	0.018	3	0.056
Smoke_Float	-2.394	0.032	3	0.053

(3) AOI Average Fixation Duration

As shown in Figure 18, experts have longer average fixation durations on the Bridge Here sign, the berthed FFG-7 bow (which is ahead of their ship), and the Instrument panel indicators for both Engines and Rudder angle; novices and experts have the same fixation duration for the Speed over Ground indicator.

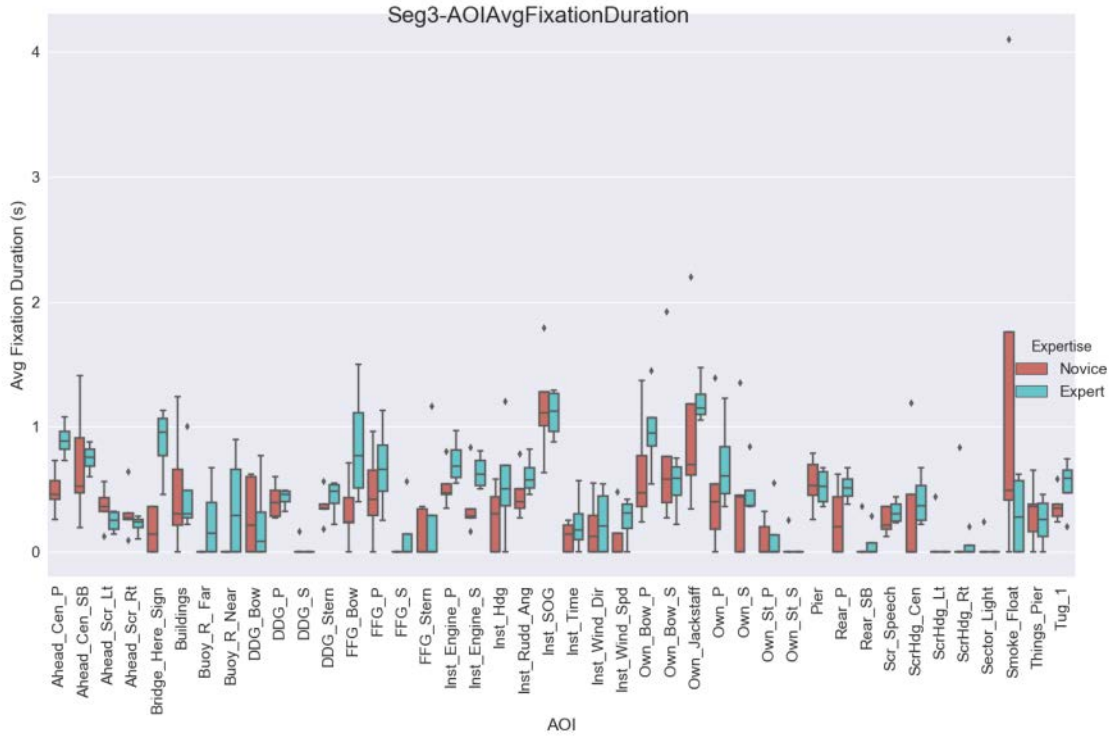


Figure 18. AOI Average Fixation Duration in Segment 3

Table 12 shows that experts have longer average fixation duration than novices on the Bridge Here sign, left half of center screen, bow of FFG-7, Instrument panel indicators and left half of stern view screen.

Table 12. T-test and MW U-test Result for Seg-3 Average Fixation Duration

AOI	T-Statistic	p-value (T-test)	MW U-Statistic	p-value (MW Test)
Bridge_Here_Sign	4.158	0.005	20	0.009
Ahead_Cen_P	3.799	0.003	19.5	0.013
FFG_Bow	1.955	0.058	17	0.056
Inst_Engine_P	1.652	0.074	17	0.056
Inst_Engine_S	1.902	0.051	16	0.089
Inst_Rudd_Ang	1.252	0.125	16	0.089
Own_Bow_P	1.203	0.134	16	0.088
Rear_P	1.936	0.051	16	0.088

b. Segment 4 Metrics

(1) AOI Fixation Count Percentage

In segment 4, there are interesting differences between experts and novices (see Figure 19). The fixation count percentages for the pier and Bridge Here sign are more than 30% for the experts, while for novices they are less than 20%. Another main area of interest for experts is the stern port side; they have more variability here, however. There are also more obvious differences in fixating on the Instrument Panel Port Engine and Starboard Engine indicators, while for the Rudder angle indicator there are no obvious differences. The experts have fixated much less than novices on the jackstaff in segment 4 and also on the smoke float and wind direction indicator on the instrument panel. Novices on average fixated on the smoke float about 1% of the time, whereas experts have never looked at it during segment 4.

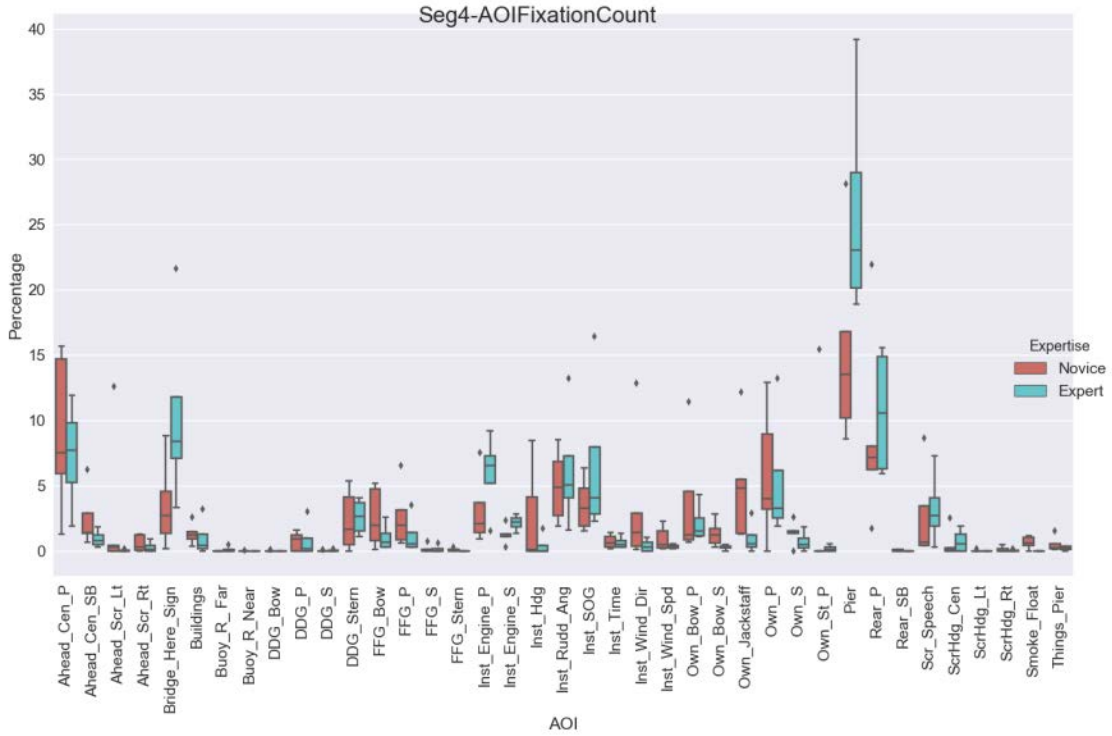


Figure 19. AOI Fixation Count Percentages in Segment 4

Table 13 shows that experts have more frequent fixations on the starboard engine indicator on the instrument panel, pier and Bridge Here sign than the novices. It also shows that experts have fewer fixations on the right half of the center screen, the stern of FFG-7, the starboard side of the bow, the jackstaff and the smoke float.

Table 13. T-test and MW U-test Result for Seg-4 Fixation Count Percentage

AOI	T-Statistic	p-value (T-test)	MW U-Statistic	p-value (MW Test)
Inst_Engine_S	1.965	0.045	18	0.033
Pier	1.842	0.058	17	0.055
Bridge_Here_Sign	1.651	0.088	16	0.089
Ahead_Cen_SB	-1.513	0.096	4	0.089
FFG_Stern	-1.675	0.084	4	0.055
Inst_Wind_Dir	-1.299	0.131	4	0.089
Own_Bow_S	-2.329	0.037	2	0.033
Own_Jackstaff	-1.929	0.057	2	0.033
Smoke_Float	-3.087	0.018	2	0.022

(2) AOI Total Fixation Duration Percentage

Figure 20 shows that in terms of fixation duration percentages, other than the pier, Bridge Here sign, stern port side view and engine indicators on the instrument panel, experts fixated for more duration on the speed indicator, a difference that was not obvious for percentage count of fixations.

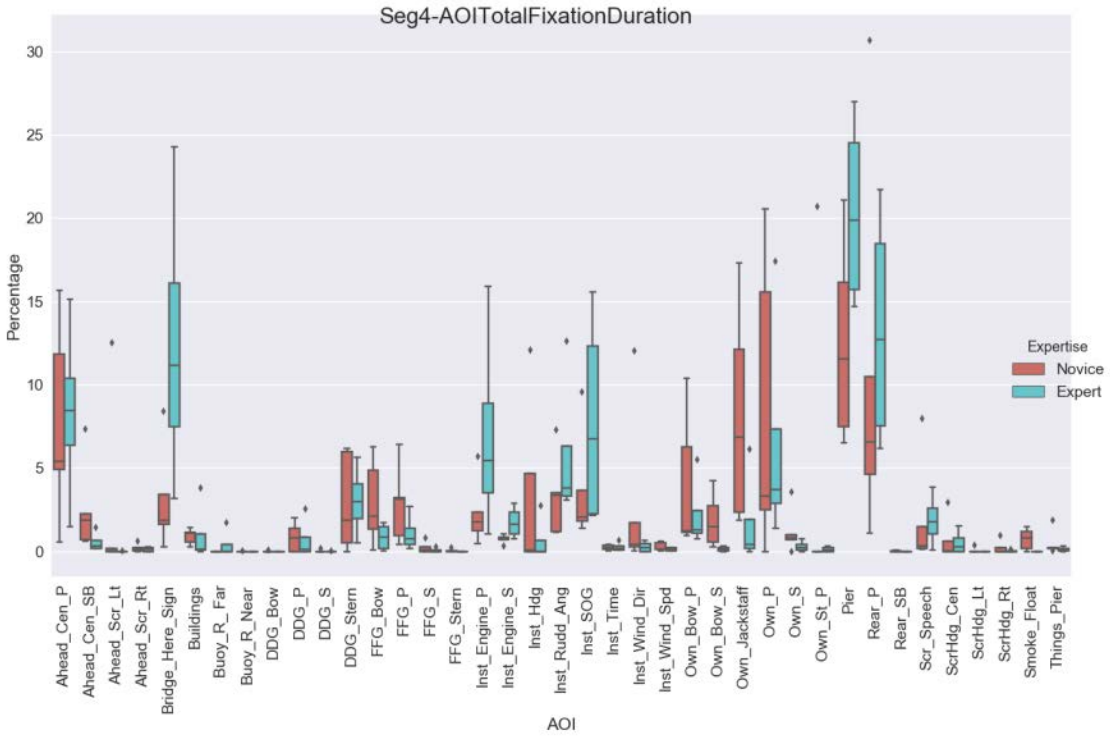


Figure 20. AOI Fixation Duration Percentages in Segment 4

Table 14 shows that experts have more fixation duration percentage on the Bridge Here Sign, starboard engine indicator on the instrument panel and pier. Experts have less total fixation duration percentage on the port side of FFG-61, right half of the center screen, jackstaff, smoke float and right side of own-ship bow.

Table 14. T-test and MW U-test Result for Seg-4 Total Fixation Duration Percentage

AOI	T-Statistic	p-value (T-test)	MW U-Statistic	p-value (MW Test)
Bridge_Here_Sign	1.993	0.062	18	0.033
Inst_Engine_S	1.976	0.066	17	0.056
Inst_SOG	1.119	0.161	16	0.089
Pier	1.937	0.048	16	0.089
FFG_P	-1.468	0.096	4	0.089
FFG_Stern	-1.327	0.128	4	0.055
Ahead_Cen_SB	-1.560	0.093	2	0.033
Own_Jackstaff	-1.928	0.052	2	0.033
Smoke_Float	-2.602	0.029	2	0.022
Own_Bow_S	-2.342	0.039	1	0.019

(3) AOI Average Fixation Duration

For average fixation duration, there are no clear differences except for a few instrument panel indicators where experts tend to have longer average fixations than novices (Figure 21).

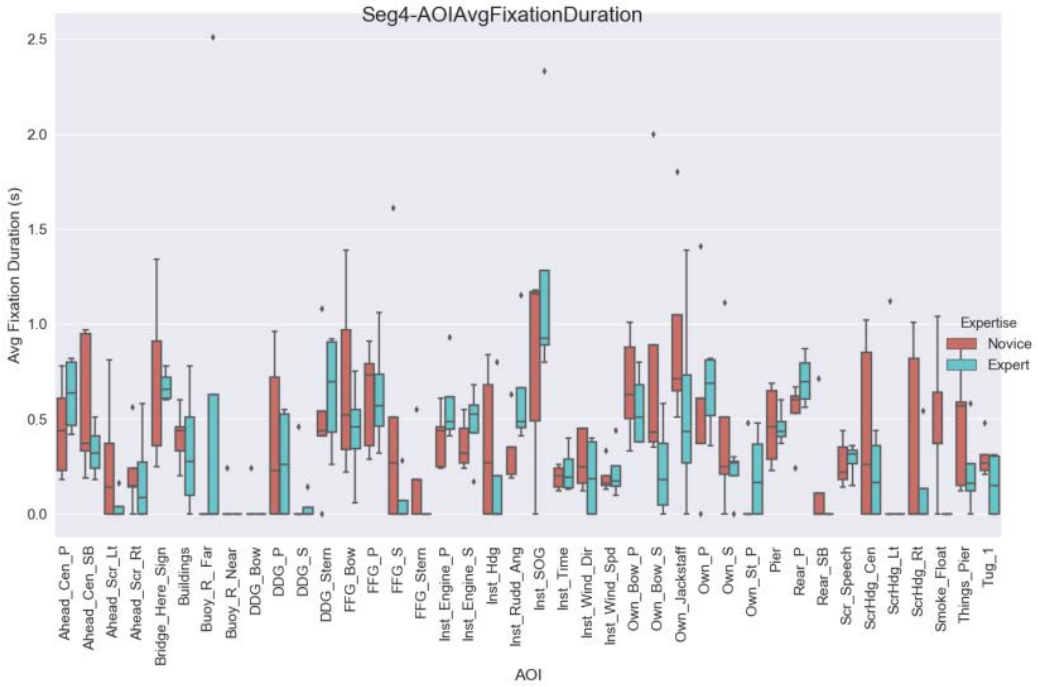


Figure 21. AOI Average Fixation Duration in Segment 4

Table 15 shows that experts do not have longer average fixation duration in any AOI, while they have less average fixation duration on the FFG-7 stern, starboard side of own-ship bow and smoke float.

Table 15. T-test and MW U-test Result for Seg-4 Average Fixation Duration

AOI	T-Statistic	p-value (T-test)	MW U-Statistic	p-value (MW Test)
Inst_Rudd_Ang	1.504	0.102	17	0.055
FFG_Stern	-1.812	0.072	4	0.054
Own_Bow_S	-1.691	0.074	3	0.056
Smoke_Float	-2.809	0.024	2	0.022

2. Eye-Scan Graphs

Eye-scan data depict where a participant looked and when, as well as their scan transitions between AOIs. Thus, eye-scan data are an important way to try to understand

participants' shiphandling strategies. Eye scan can be looked at in a number of ways. In this study, we built two kinds of visualizations: network graph and timeline charts.

Network graphs are used here to visualize the eye-scan fixation transitions between AOIs. These graphs consist of the nodes and edges:

- The nodes show the fixated AOIs.
- The size of node indicates the number of adjacent nodes (AOIs).
- The edges between nodes indicate the transitions between nodes. The thicker side of edge indicates the arrow. Where both sides are thicker, it means a two-sided edge.
- The edge color indicates the number of transitions between AOIs. The reference is given in the figure legends.

The timeline chart is used to depict AOIs fixated over time. It shows the focused areas of participants over time. Network and timeline graphs were created for each participant. In the following paragraphs, we present scan graphs of an expert and a novice to show comparison of their fixation patterns during segment 3 and segment 4. The scan graphs for the remaining participants are located in Appendix D.

a. Eye Scans in Segment 3

The eye-scan comparison is given for P-7-7 (expert) and I-7-4 (novice). P-7-7 was chosen as he showed the best performance even among instructors and I-7-4 was chosen as a random student who was able to complete the exercise. Segment 3 is detailed in Figures 22 and 23.

(1) P-7-7 Segment 3

The network graph shows that the expert has a large number of transitions between the jackstaff (middle of the center screen), left half of the center screen and left side of own-ship's bow. The expert's fixation has also transitioned between own-ship's port side and jackstaff, and between the pier and general lookout in the left half of the center screen. This shows that he has mostly fixated on the left half of the center screen and on the stern view screen. While looking at the time scale chart, we can see that from close to eight minutes into segment 3, there is significant shift to looking at the pier, own-

ship’s port side and port engine indicator on the instrument panel. There are also several fixations on the stern of the berthed DDG and own-ship stern during later times.

(2) I-7-4 Segment 3

As compared to the expert, the novice is mostly focused on the instrument panel’s wind direction and rudder angle indicators and had a lot of transitions in between these indicators. This is a place where the expert hardly ever looked during segment 3. Additionally, the novice is focused on the pier right from the beginning of segment 3 and this frequency lessened towards the end of segment 3, in direct contrast to the expert’s scan. The novice also has fixated more on the port side of the DDG berthed at the pier than the stern, where the expert was more fixated. There are many transitions to the speech screen from different AOIs, which indicates he is repeatedly confirming his voice commands written as text on the screen—something that is useless during this exercise. Also, in his network graph, there are quite a number of nodes (AOIs) with large size. This result means that he had quite a scattered pattern in comparison with the expert, who had a much tighter scan pattern.

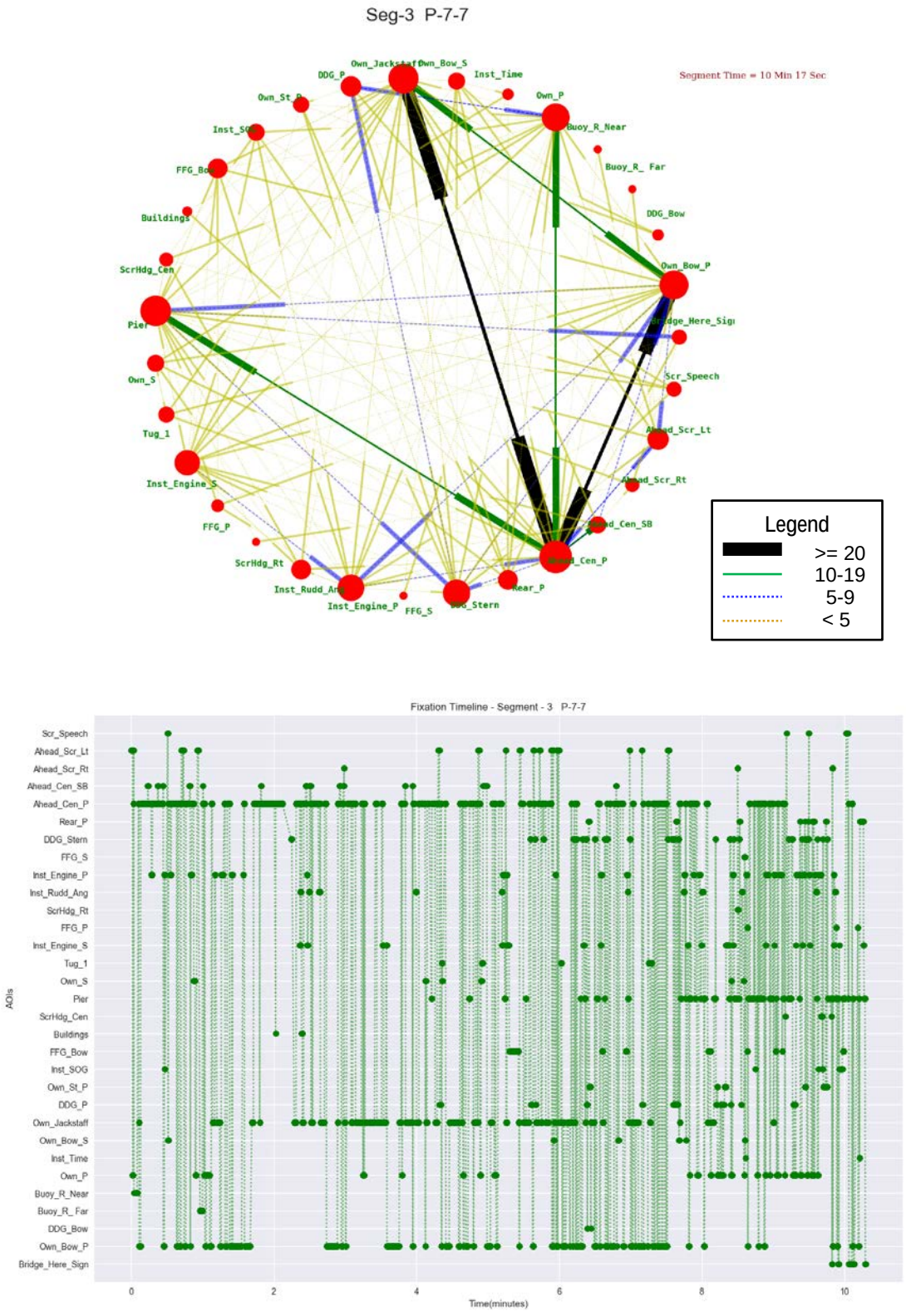


Figure 22. Eye-Scan Network and Timeline Graphs of P-7-7 in Segment 3

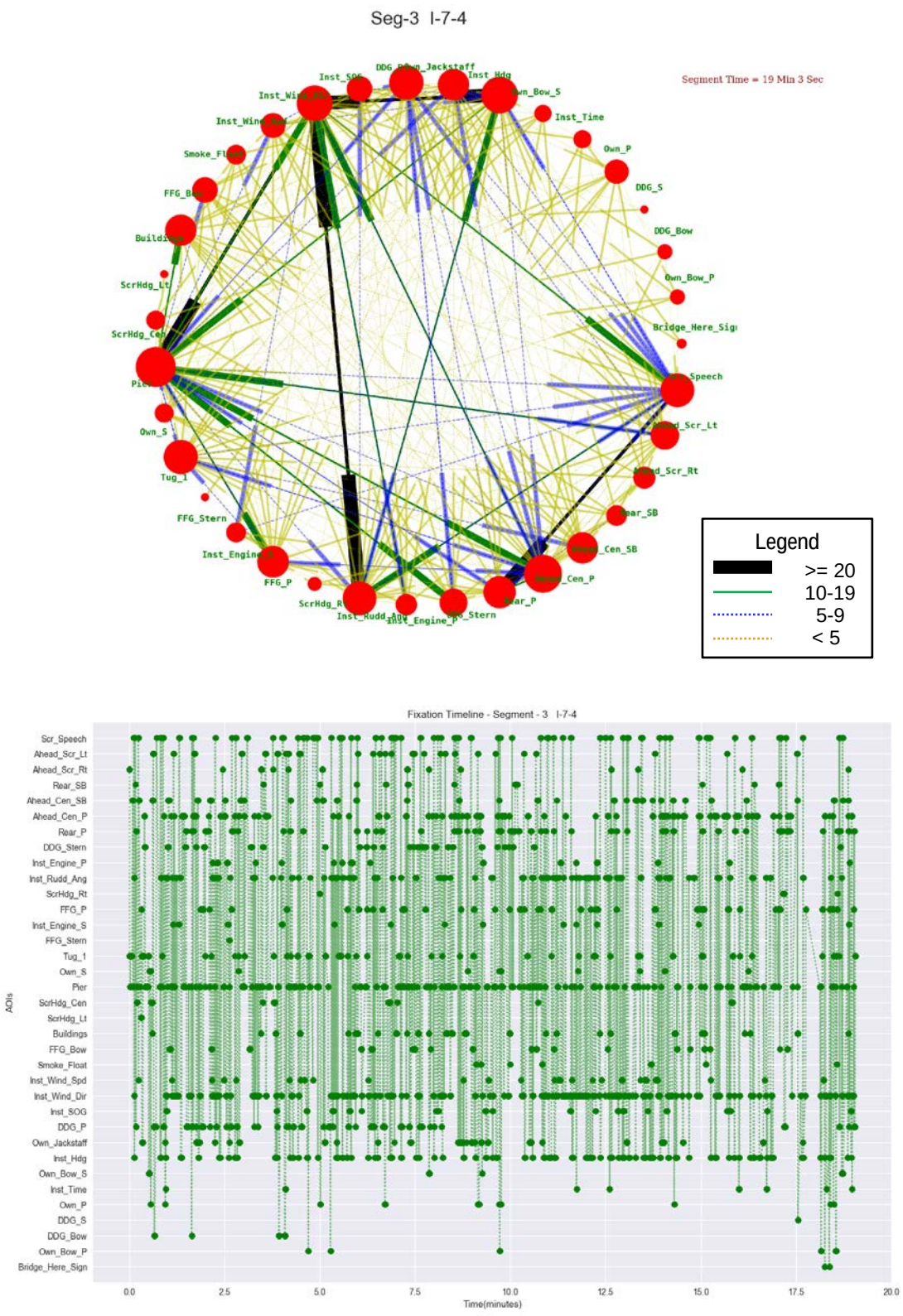


Figure 23. Eye-Scan Network and Timeline Graphs of I-7-4 in Segment 3

b. Eye Scans in Segments-4

The eye-scan comparison for P-7-7 (expert) and I-7-4 (novice) during segment 4 is detailed in Figures 24 and 25.

(1) P-7-7 Segment 4

In segment 4, the expert has shown a tight scan pattern between the Bridge Here sign, pier and instrument panel port engine indicator. There is also a pattern between the Bridge Here sign, pier and speed over ground indicator. On the timeline graph, an interesting shift towards the end of the exercise is from the Bridge Here sign to own-ship port side. This shift indicates that after aligning with the Bridge Here sign, the shiphandler is now concentrating on safe contact of the port side with the pier.

(2) I-7-4 Segment 4

In comparison to the expert, this novice has not shown any specific or tight scan pattern in segment 4. There are a lot of transitions from heading indicator and rudder angle indicators to wind direction indicators on the same screen, showing the novice's focus to get information from the instrument. The expert, on the other hand, has been more concentrated on visual cues from the outside world. The novice has shown very little concentration on the Bridge Here sign. Rather, most of the time he fixated on other places on the pier. The number of connections are also large for a number of nodes.

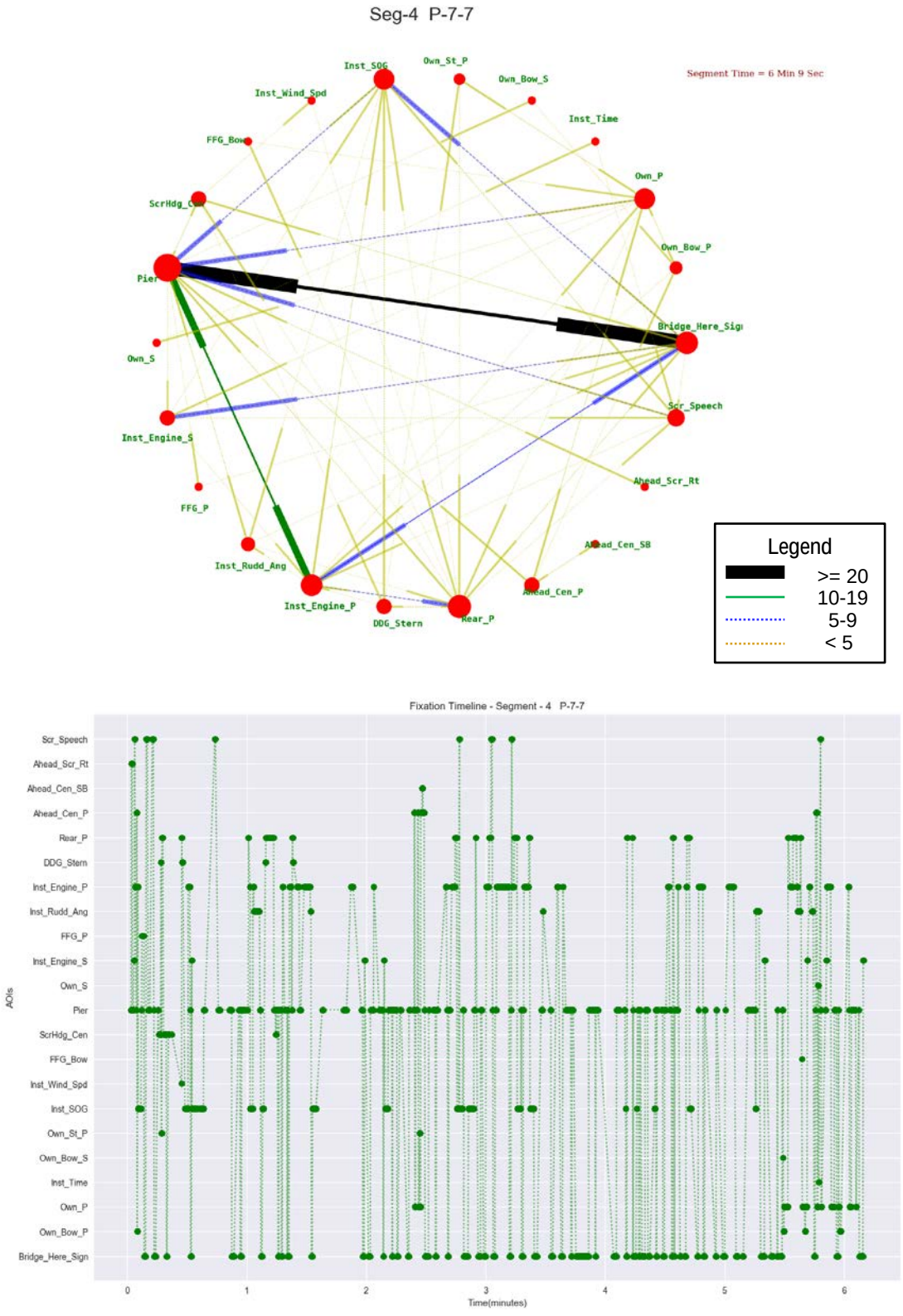


Figure 24. Eye-Scan Network and Timeline Graphs of P-7-7 in Segment 4

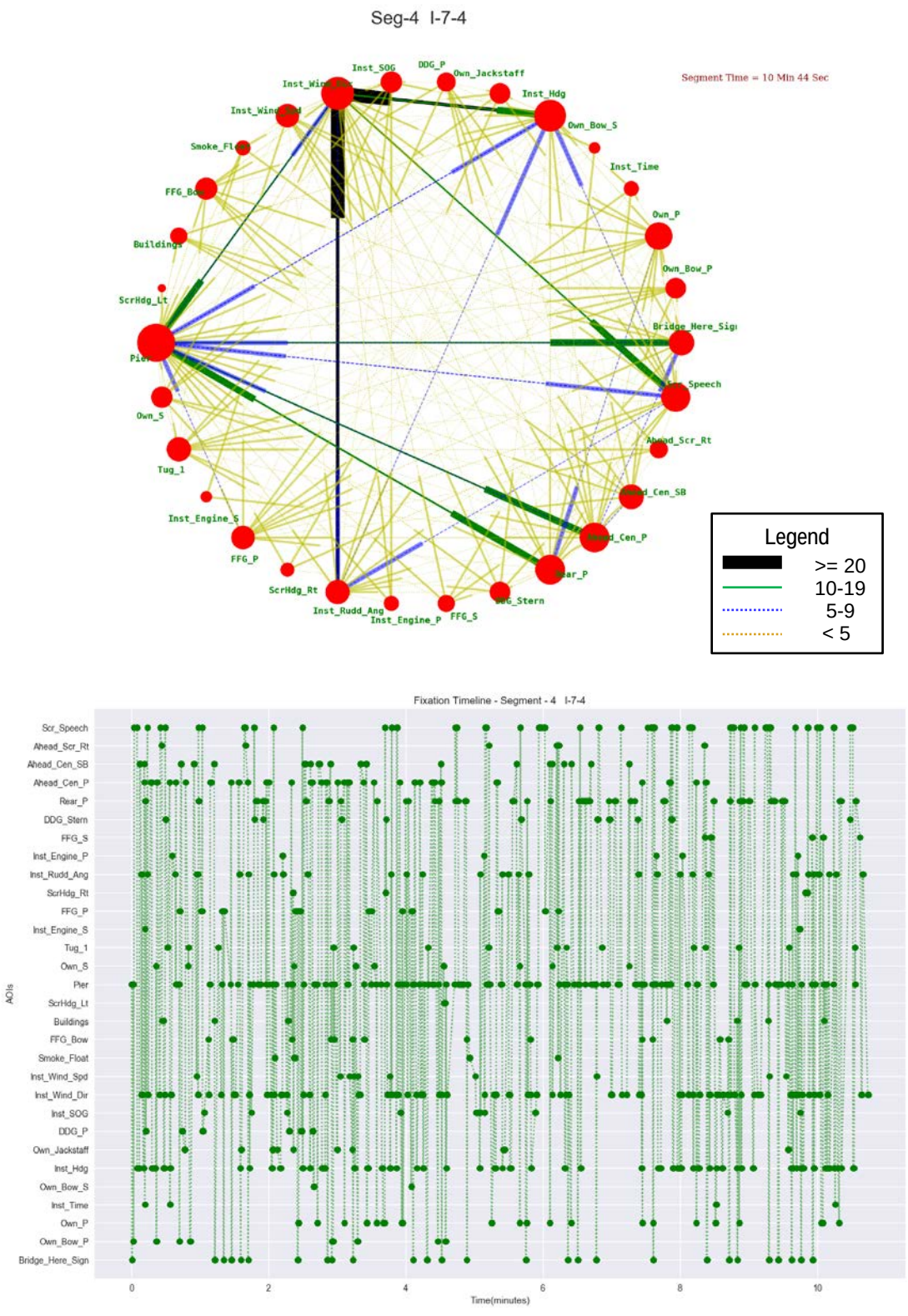


Figure 25. Eye-Scan Network and Timeline Graphs of I-7-4 in Segment 4

c. Summary of Eye-Scan Graphs

There are obvious differences between novice and expert patterns and fixation area. If we couple this with their performance by visualizing the ship's location, heading and speed at these times, then the experts fixated on the key areas at the right time. These visual strategies are absent from the novices' eye tracking visualization. Hence, we can safely deduce that the experts' Attention-allocation patterns aid their shiphandling performance.

C. PERFORMANCE ANALYSIS FROM SIMULATION DATA

The aim of this study is to improve the feedback of the ITS system for trainee shiphandlers and instructors. While eye-tracking data provides insights into the cognitive strategies being used by the trainees, their performance in the simulation is measured by the ship's parameters being maintained during the exercise. This section describes a method devised to score the performance of shiphandlers using simulation data.

Simulation logs stores ship's parameters at a frequency of 1 Hz. It includes speeds, position, heading, timestamp, acceleration, etc. From these parameters, distance and bearings of the ship from the Bridge Here sign were calculated and data were truncated for segments-3 and 4 only to gauge the performance of novices. As per the sponsors' guidance, a number of measures of performance (MOPs) were selected to predict the performance. These MOPs are described in the following paragraphs.

1. Measures of Performance

The distance of the ship from the Bridge Here sign at the start of segment 3 is approximately 800 yards. Within this distance, at every position starting from 800 yards until 50 yards from the Bridge Here sign, we have considered four parameters to decide upon the driver's performance. These are (Figure 26):

1. Time to reach that position
2. Bearing of ship from pier
3. Orientation of ship (i.e., its current heading with respect to north)
4. Longitudinal speed of the ship at current position

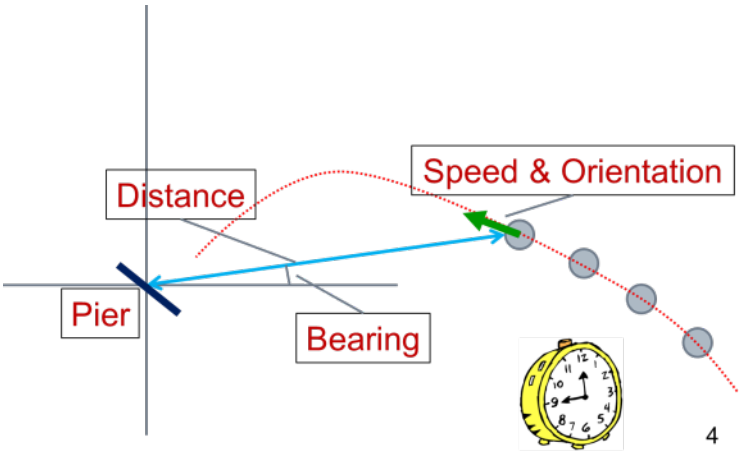


Figure 26. Parameters Used for Evaluating Shiphandler's Performance

2. Criteria and Evaluation

To develop the criteria for evaluating the novices' data, we used the four available experts' recordings. We aggregated the data of four MOPs for every 10 yards starting from 800 to 50 yards. We calculated their mean and standard deviations within this 10-yard block across all experts. This data served as the criteria for evaluating every position in the novice paths. For each MOP, every data point was scored between 0.0 and 1.0 based upon its difference from aggregated experts' criteria.

3. Analytical Hierarchy Process

After scoring each point for four MOPs (one score for each MOP), a single performance score was awarded to each point based upon weights calculated from AHP. The AHP matrix was formulated based upon the subject knowledge provided by the sponsor (Table 16).

Table 16. Weights Matrix for AHP

	Brg	Orient	Time	Speed
Brg	1	3	1	1/7
Orient	1/3	1	1/3	1/7
Time	3	3	1	1/3
Speed	7	7	3	1

The maximum eigenvalue obtained from the matrix is λ_{max} = 4.49 (10% tolerance value is 4.27). Hence, the weights calculated from normalized eigenvectors were as follows:

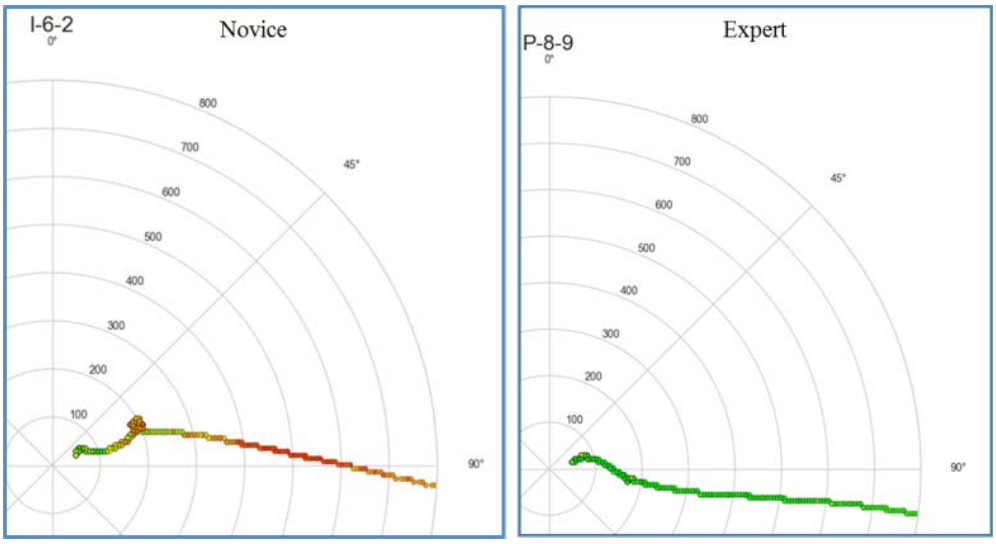
- Bearing: 0.137
- Orientation: 0.058
- Time: 0.222
- Speed: 0.582

The final performance score for each data point was thus calculated as:

0.137 * BrgScore + 0.058 * OrientScore + 0.222 * TimeScore + 0.582 * Speed Score

4. Expert vs. Novice Performance

The scores obtained from the data were plotted in a polar graph for each participant, as shown in Figure 27 for an expert and a novice. Each point represents a data point recorded at 1-second intervals. The center of concentric circles indicates the Bridge Here Sign on the pier. The circles indicate distance in yards from the Bridge Here Sign, while angle indicates the azimuth (bearing) from the Bridge Here Sign with respect to North. The initial position of the ship is at the extreme circle (i.e., 800 yards) and the final position is close to the center. The best performance score is 1.0 displayed in green, the worst score is 0.0 displayed in red, and the middle score of 0.5 is displayed in yellow. All the colors in between are obtained by linear interpolation between these three colors. Figure 27 depicts the difference between an expert and novice's performance.



The performance of the expert indicates his score remained closer to 1 at every recorded location, while the novice's score fluctuated between to the extreme values of 0 and 1.

Figure 27. Comparison of Expert and Novice Performance

5. Novices' Performance Evaluation

The evaluated performance of novices is shown in Figure 28 as per the described algorithm. I-6-2's graph shows yellow/red colors in the beginning because instead of approaching the pier from the last buoys, he continued heading straight in the last maintained direction, which was contrary to the experts' path. Close to 350 yards, he started to close in but struggled at close to 200 yards to achieve lateral motion of the ship. Finally, he managed to land the ship along the pier, as illustrated by the green points close to the center. I-6-3 performed well in the beginning until around 130 yards, at which point he struggled to close in, taking a lot of time. He eventually reached the objective. I-7-4 struggled in the middle of segment 3 and also close to 200 yards from the final destination. I-7-5 did well until 200 yards but after that, instead of closing in, he was constantly going away from the pier. He walked his ship upwind with little-to-no fore/aft motion, which is more difficult than walking it downwind. However, successfully reach the pier required walking it downwind and he could not adapt. Perhaps, he needs more coaching in using the engines and rudders to achieve the lateral movement of the ship, which is a skill that trainees get to practice and experiment with.

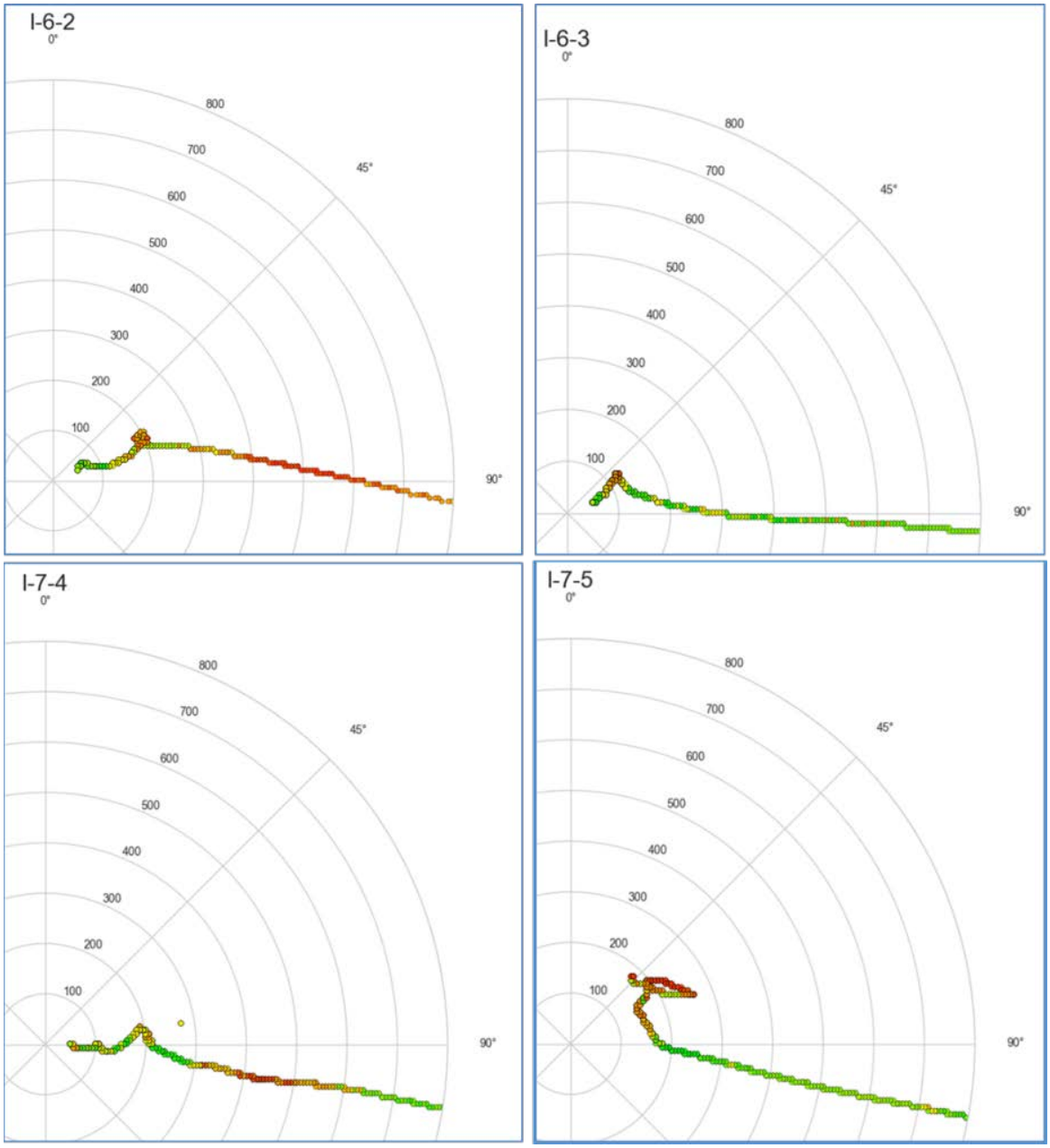


Figure 28. Performance Evaluation of Novices

D. POST TASK SURVEYS

Participants' responses to the post-task survey regarding their shiphandling strategies, the visual cues they used, and self-evaluation of their performance at each segment of the exercise are listed below. Interestingly, most of the novices evaluated their performance very highly. Two exceptions were participants I-7-5 and I-8-8. Participant I-7-5 evaluated himself below par in all the stages and participant I-8-8 failed himself during the last stage, as he crashed the ship on the pier while trying to adjust the front screen view on the simulator.

The feedback from experts' and novices' use of visual cues shows more consistency among experts than novices. During stage-1, all the experts have listed the sector light, jackstaff and buoys as their visual cues. In stage-2, the common visual cues include the Bridge Here sign and ships berthed at the pier. In stage-3, the common visual cue used is the Bridge Here Sign. The analyses of eye-tracking data confirm the use of visual cues by experts as reported in post-task surveys, however, eye tracking reveals much more details then what they just reported. This suggests that, for experts, to focus on some AOIs in a situation becomes automatic with experience and they might not be able to exactly report it in detail. The novices on the other hand have shown a mixed behavior. Eye-tracking data shows their use of reported visual cues, however, there are cues reported in post task surveys that were hardly used by them during the exercise as shown by eye-tracking data.

Experts also have listed more common shiphandling strategies they used during the simulation exercise than novices. The post-task feedback was collected for three different phases of the exercise. It included approaching the berthing location, lining up with the pier and then closing in to the pier. Responses are tabulated in Tables 17 through 19.

Table 17. Self-Evaluation, Strategies and Visual Cues Used by Participants from Start Point to Approaching the Berth

ID	Self-Evaluation	Strategies	Visual Cues
I-6-2	90%	-Used range lights before turn to stay in channel -Used sector light to stay in safe water -Kept green buoys starboard -Slowed to 5 kts for tug makeup and 3 kts to approach pier	-Range lights -Sector lights -Buoys -Pier -Tug
I-6-3	90%	-Used lot of estimation of swing of ship without using VMS -When pulling, used normal technique of turning slightly to pier -Performing a twist with port engine back 22% and SB engine ahead 1/3 -Let the environmentals ease the ship to the pier	-Horizon reference to ship's aft stack and jackstaff -Smoke float -Rate of close between specific points on ship
I-7-4	90%	-Used ranges to include distances -Knew buoy layout to judge turn and speed checks	-Used relative motion in relation to stationary objects -Lineup lights
I-7-5	20%		-Buoys
I-8-8	90%	-Planned -Know how the ship moves	-Vector Lights -Speed
III-8-10	90%	-Used buoys -Sector light as reference -Try to maintain visually middle channel	-Sector light -Jackstaff -Buoys
P-7-7	85%	-Lineup with Sector Light -Min use of rudder to maintain course	-Sector Light -Buoys -FFG at Dock
P-8-9	100%	-Rudder in Hand -Slow & Steady -Small Adjustment -Keeping her between Buoys	-Range Lights -Sector Light -Position of ships on pier -Smoke signal -Items on Pier -Closure Rate to Pier -Jackstaff

Table 18. Self-Evaluation, Strategies and Visual Cues Used by Participants for Getting the Ship Lined up with the Pier

ID	Self-Evaluation	Strategies	Visual Cues
I-6-2	90%	-Lined up Bridge Here Sign to 270R of own ship -Adjusted fwd/aft line using backing engine speed changes <5% pitch for <0.2kts fwd/aft motion -Ensured shallow angle of approach (<8°) to ensure relative visuals remain accurate within 10°	-Bridge Here Sign -SOG Indicators -Pier Heading 302T to check angle of approach <8°
I-6-3	90%	-Focused on landing the bow and stern -Getting the ship on pier heading -Speed down to zero -Maintaining control with twisting the ship	-Used lot of jackstaff and stern tracking with horizon to determine closure rate of the ship
I-7-4	70%	-Rudder and tugs to fine tune -I misjudged the effects of environmentals so I had to readjust by overpowering with my tug and engines	-Used relative motion to near -Pointed my bow to the FFG
I-7-5	20%	-Intended to use port low speed twist counteracting the environmentals	-Bridge Here Sign -Other ships on Pier
I-8-8	90%	-Speed Control -Bridge Here Sign 270° right -Tug	-Control IRT factors
III-8-10	90%	-Tried to keep the jackstaff to the right of Bridge Here Sign -Keep good momentum past the DDG	-DDG on Pier -Bridge Here Sign -FFG on Pier
P-7-7	85%	-Lineup so environmentals got me to the pier	-Bridge Here Sign -Stern of DDG at Dock
P-8-9	100%	-Small Bells -Rudder in Hand -Small Adjustment -Slow & Steady	-Bridge Here Sign -Smoke signal -Items on Pier -Jackstaff

Table 19. Self-Evaluation, Strategies and Visual Cues Used by Participants for Bringing Ship Closer to the Pier

ID	Self-Evaluation	Strategies	Visual Cues
I-6-2	85%	-Setup low speed twist to allow on-setting environmentals to bring the ship to pier -Attempted to adjust one variable at a time, wait to see effect, then make another adjustment Used backing engines to monitor/change speed Used rudder to check stern movement Used tug to check bow movement	-Stern closure rate in reference to aft ship -Bow closure rate in reference to fwd ship/fixed pier objects Pier heading of 302T to gauge angle of approach -ROT indicator/closure visuals above to determine when to employ/stop tug
I-6-3	95%	-I utilized the tug to slowly let the bow land on the pier and the rudder to slowly let the stern land on the pier	-Looked at bow in relation to the pier -Pier heading -Stern in relation to the pier
I-7-4	95%	-After misjudging environmentals, I utilized all resources: distance, bearing, tugs, engines	-DDG closing rate -Stern to pier ratio gave me an idea of how the environmentals play
I-7-5	0%	-Low speed twist	-Bridge Here Sign -Unofficial ranges on the pier -Bow and stern movements
I-8-8	15%	-Speed -Bridge Here Sign -Know the Simulator	-The Ship -The Pier
III-8-10	75%	-Make sure I was pier heading before touching down -Tried to get bridge as close as possible to Bridge Here Sign	-Bridge Here Sign -DDG on Pier -FFG on Pier
P-7-7	85%	-Check speed with engines -Used rudder and tug to control heading	-Bridge Here Sign
P-8-9	100%	-Slow & Steady -Rudder in Hand -Minimal engine orders	-Bridge Here Sign -Relative motion with ships already moored -Jackstaff

V. CONCLUSION AND DISCUSSION

This research focused on finding the differences in gaze parameters, scan patterns and exercise performance between expert and novice shiphandlers in a challenging mooring scenario. It is a stepping stone towards incorporating the accurate cognitive state in the expert model of the ITS. An improved ITS is expected to give better feedback to the shiphandlers during simulation training. The insights gained from this study may help shiphandling instructors to better teach visual scanning strategies to the students.

A. SUMMARY OF RESULTS

The positional and time information from ship data in our study indicate differences between experts and novices, especially as the ship gets closer to the pier. The eye gaze fixation metrics analyses, like the percentage of fixation duration in AOIs, also indicate expertise differences in all segments of the exercise, especially during segments 3 and 4. The tighter scan pattern of experts marked by fewer visited AOIs and strong network connections between even a fewer number of AOIs during segments 3 and 4 further distinguishes experts from novices. The time distribution of experts' fixation in AOIs indicates clear shift of focus from certain AOIs to others as the situation progresses, while such pattern is not obvious in novices' distribution. Because we had a very small sample size in our study, we did not emphasize statistical significance in the results. The results, however, support the domain knowledge provided by the sponsor and are consistent with previous studies examining expertise differences in visual scan patterns.

B. IMPLICATIONS OF RESULTS

The results obtained in our study can have important implications towards the development of the ITS system and also in enhancing the training value of shiphandling. These implications are described in the following paragraphs.

1. Results for Classroom Teaching

The information gained from this study can be used in teaching optimal attention allocation techniques to the trainee shiphandlers during various phases of mooring

exercises. We found that the experts focused on specific controls—the rudder and the engines, along with the tug boat—during the last phases of the exercise. These results can be used to better teach shiphandlers when and how to employ these controls and their impact on lateral movement of a ship.

2. Update the Expert Model in ITS

ITS has two components: an intelligent tutor and an expert model (Kirschenbaum et al., 2010). The latter component is a cognitive model that represents an expert's performance for various shiphandling tasks. The visual cues used by experts are established through interviews with subject matter experts and head-tracking while conducting simulation exercises. These techniques are subjective and therefore not completely accurate. In contrast, eye tracking provides more accurate quantitative data on the AOIs fixated by experts in time and space. The scan patterns and focus shifts can also be determined from the raw eye-tracking data. The expert model in ITS can thus be updated with this accurate information that may result in better feedback to students from ITS.

3. Live Multi-student Performance Display for Instructors

The algorithm developed in this study to analyze the novice performance based upon experts' simulation data weighed as per analytical hierarchical process can be a very powerful tool for tracking the live performance of students. The data from a number of COVE simulators can be integrated and analyzed on one system with multiple screens, as shown in Figure 29. The performance parameters of the student may be compared with the experts' data in that situation and a score may be displayed in color-coded form to indicate the performance of the student. This setup may be installed in instructors' rooms where they can monitor performance of all the students doing simulation and identify those in need of direct supervision. This setup may help to reduce the burden on instructors by attending students only when it is required.



Figure 29. Impression of COVE-ITS Live Performance Monitoring System

4. Debrief System

The data from eye-tracking hardware, once implemented in the COVE system, can be used in conjunction with a performance evaluation system using MoPs and AHP technique described above. This system can be developed into a very effective debrief method where instructors can show the students their cognitive states at various phases of exercise along with their performance and can coach them on their cognitive skills and ideal cognitive behavior. Such a system can be very valuable for students struggling with skill development in shiphandling.

C. LIMITATIONS

Our study's main limitations were small sample size of participants and data collection problems with eye trackers.

1. Small Sample Size

A big limitation is the small sample size. Out of 11 volunteers, two had to be excluded due to ETG wearing/calibration problems. One had to be excluded after recording due to not falling in the category of either novice or expert in our scenario. One participant's simulation data was not recorded, while another participant's voice data got

corrupted. Additionally, we lost a one-and-a-half days of data collection time due to a snowstorm and base closure. In the end, we had eye-tracking data of four experts and five novices but full simulation/voice data of three experts and four novices only. Due to this small sample size, the statistical significance of observed differences between the expert and novice groups could not be established. Further, because a shiphandling task may be performed in many different correct ways, we may not have captured a full range of expert behavior.

2. Eye-Tracking Data Collection Problems

The problems encountered during our study were mainly related to the eye tracking glasses. These problems are highlighted in the following paragraphs:

a. Eye-Tracking Quality

The data being recorded with Tobii glasses has a quality indicator in Tobii Controller Software. In our data collection, this indicator never met the good data collection guidelines given in the software manual. This problem persisted throughout the study. Figure 30 shows the exact guidelines in the manual on the software's "Track Status" tool. It states that the tracker records with the highest quality when the eye circles have a white outline, the yellow circle touches the white outline, with a large dark green area and a pupil dot within the green area. In our experience, however, we never got this situation. The manual says that if desired quality is not being achieved, then we should change the nose piece and tighten the head band. All these procedures were performed, but even the best achieved results were showing small green circles and slightly bigger yellow circles. These results were also not for every participant; it was observed that lowest quality was achieved with people having deep-set eyes. Furthermore, in almost every case, one eye circle was consistently smaller than the other eye circle. This discrepancy was observed even in the case of the second pair sent by the manufacturer's support office. During the recording, it was observed that circle size fluctuates and, many times, one eye would become red and lose its green and yellow circles.

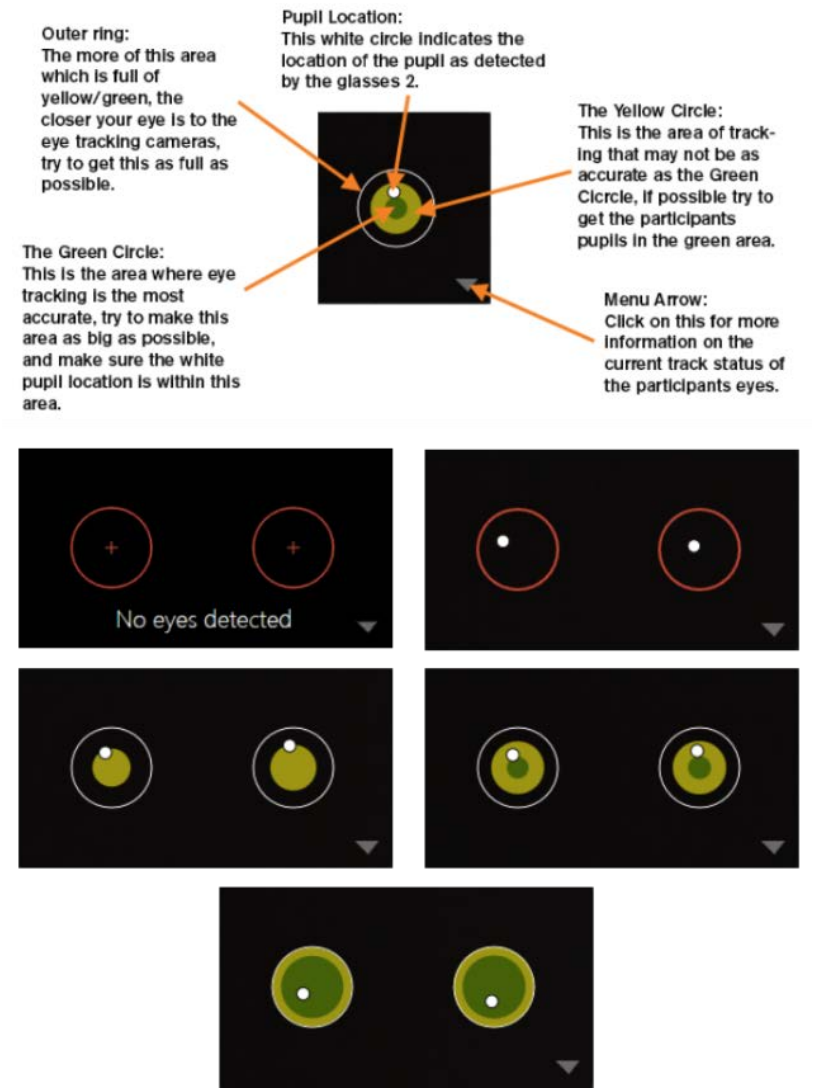


Figure 30. Track Status Tool Guidelines from Tobii Pro Glasses 2 User Manual

b. Validating the Object Tracked by Subject's Eyes

Another issue was with the validation of objects being tracked by the subject. Validation was done by asking the subject to focus on a specific object on the screen and cross-checking it with the circle representing their eye tracking position on the controller software. We often observed that the eye tracking circle was consistently showing a lower position than where the subject was actually looking. This problem persisted and had to be adjusted by manual mapping of gaze data on AOI chart during data preparation.

c. Unserviceability of ETG during Initial Study

During our initial visit to SWOS for coordination and initial study, we tested the ETG with our study team. Members of the study team ran some scenarios on COVE wearing ETG. During trials, the Tobii controller software indicated the left eye signal was of poorer quality than the right eye. Although we faced calibration issues with some people, as we continued to use the glasses, the problem kept increasing. Ultimately, the glasses stopped recognizing the left eye and calibration started to fail for everyone. We contacted the manufacturer support in North America, who tried to fix the issue online by upgrading the firmware, but this did not resolve the problem. We ultimately had to send the glasses back to them to fix the issue; meanwhile, they facilitated our research by loaning us another ETG to continue our study. Figure 31 gives an example of these problems.



Figure 31. Screenshot of Eyes as Recognized from Glasses Controller Software

Another problem faced with ETG was the issue of battery time. Fully charged batteries would initially show recording time of more than an hour, but would deplete within 20–25 minutes. This problem was observed only with our own newly purchased glasses and not with the ones sent by the manufacturer's support office.

3. Data Loss in Simulation and Voice Logs

This issue occurred in the case of one participant only. Due to some unknown problem, the simulation exercise data and voice logs recorded by the COVE system were corrupted for one of the participants and could not be obtained. For another participant, only the voice log was corrupted but simulation data was intact.

D. RECOMMENDATIONS ON FUTURE WORK

This work can be applied to any military task where use of visual cues is necessary and more experience results in efficient scan patterns. Similar studies may be carried out in areas like aviation, radars / sensors operations, submarine handling, air traffic control, weapons firing etc. Specifically, this is a pilot study towards improving the ITS using cognitive state of shiphandlers from eye-tracking data. In the future, this work can be extended in the following manner:

1. Further work may be continued by collecting and analyzing more experts' and novices' data to establish the significance of the results found on differences in fixation metrics and scan patterns.

2. A feasibility study may be carried out to fit eye tracking sensors into the head-mounted display of the COVE-1 system, as it is widely used in both the Basic Division Officers Course and ADOC. Furthermore, feasibility for incorporating eye tracking glasses in the COVE-3 system on a permanent basis may be conducted.

3. Research may be done to automatically identify the fixated object of interest and to integrate live gaze information from the eye tracking devices into the ITS to ascertain the cognitive state of shiphandlers during the exercise. This could be done either by computer vision techniques automatically identifying the objects being observed by the user or by modifying the visualization code to allow the system to identify the object at the pixel the user is looking at on the screen.

4. Machine learning techniques may be used after collecting more experts' data to build models of ideal cognitive strategies and performance parameters. These models can then be used to analyze and score the simulation exercises and to inform the instructor of students' weaknesses.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. DEMOGRAPHICS SURVEY FORM

Demographics Survey

Date: _____ **Participant's ID :** _____

1. **Age:** _____
2. **Years commissioned service:** _____
3. **Gender:** Male Female
4. **Commissioning source:**

   Maritime Academy / NROTC / OCS / USNA
5. **Highest tour completed:** 1st DIVO 2nd DIVO DH XO CO
6. **FOR ADOC STUDENTS ONLY: Are you officer of the deck qualified?**

   Yes / No a. If so, how many months have you been qualified? _____
7. **Time at Sea:** In the table below, please estimate how many months you served AS A COMMISSIONED OFFICER on each of the classes below. For each class, indicate both total time aboard as well as time spent routinely standing bridge watches. Also, estimate how many times you served as conning officer/OOD or XO/CO during a sea and anchor transit.

Ship Class	Months Aboard	Months U/W Watch on Bridge	Sea and Anchor Transits as conning officer/ OOD/XO/CO
CV/CVN/LHA/LHD			
Other amphibious ship			
CG-47			
DDG-51			
FFG-7			
LCS			
MCM/PC			
Other			

8. Do you have prior enlisted service? Yes / No (Skip to 9)

   a. Years enlisted: _____

   b. Highest rate Achieved: _____

   c. Years of enlisted sea duty: _____

   d. Did you serve as coxswain? Yes / No
9. Have you used a simulator for shiphandling training?

   Yes / No a. If yes, how many times? _____
10. On a scale of "1" (strong disagreement) to "7" (strong agreement), rate your feelings about the statement, "Simulation is an effective learning tool for shiphandling."

    Strongly Disagree Strongly Agree

    1 2 3 4 5 6 7

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. POST-TASK SURVEY FORM

Post Task Survey



Date : _____

Participant's ID : _____

Stage-I: Approaching the berth (From the beginning until lining up the ship with berth)	
1.	Draw a vertical line indicating how well you think you did on this task: |----------| Failed Best
2.	List the strategies that you used:
3.	List the visual cues that you used:
Stage-II: Getting the ship lined up at the with berthing location	
1.	Draw a vertical line indicating how well you think you did on this task: |----------| Failed Best
2.	List the strategies that you used:
3.	List the visual cues that you used:

Stage-III: Bringing the ship close to the pier	
1.	Draw a vertical line indicating how well you think you did on this task: |----------| Failed Best
2.	List the strategies that you used:
3.	List the visual cues that you used:

APPENDIX C. EYE-TRACKING METRICS SEGMENTS- 1 AND 2

Segment 1 Metrics

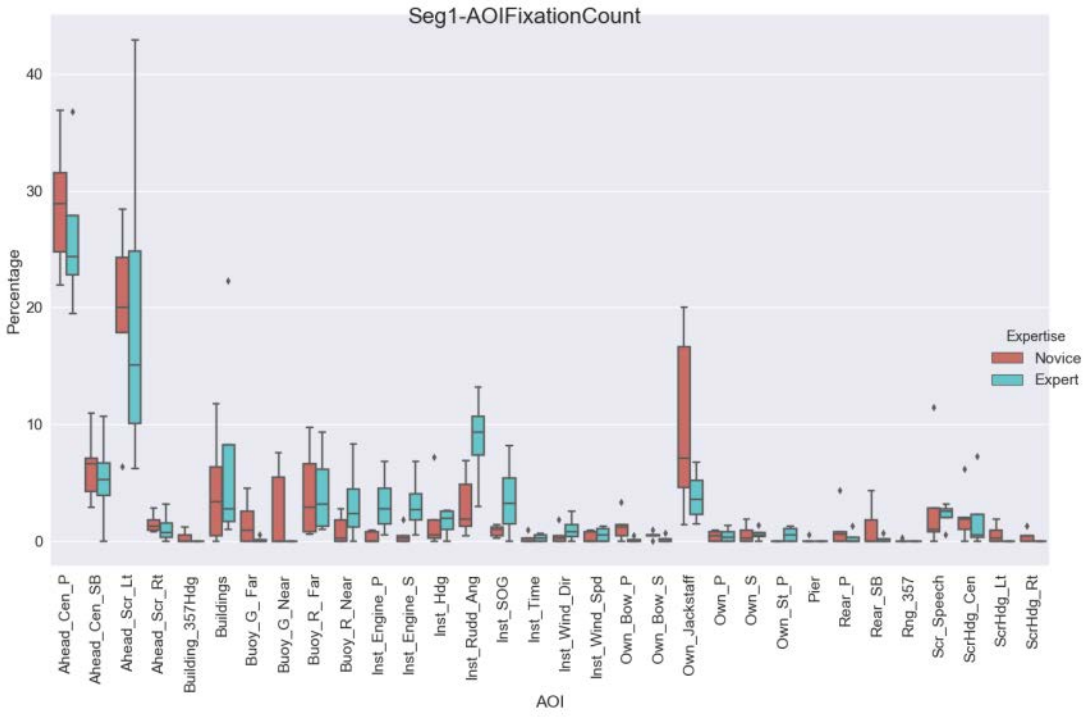


Figure 32. AOI Fixation Count Percentages in Segment 1

Table 20. Mann-Whitney U-test Result for Seg-1 Fixation Count Percentage

AOI	MW U-Statistic	p-value
Inst_Engine_S	19	0.018
Inst_Engine_P	18	0.031
Inst_Rudd_Ang	18	0.033
Own_St_P	15	0.066
ScrHdg_Lt	4	0.055
ScrHdg_Rt	4	0.055
Own_Bow_P	2.5	0.037

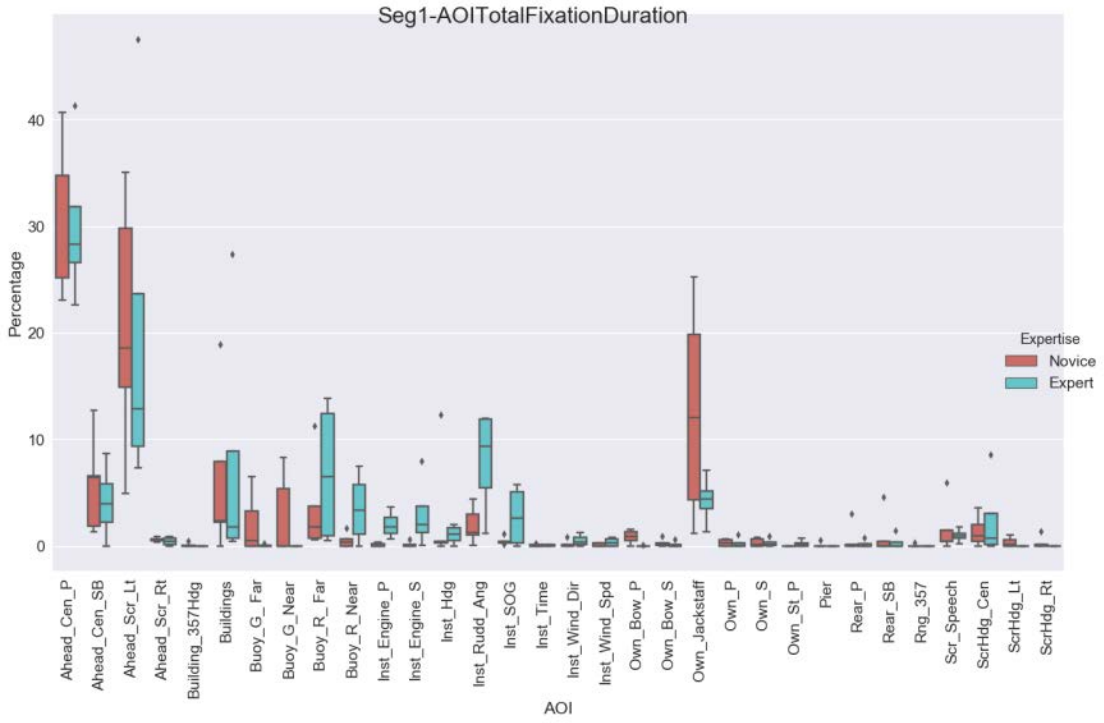


Figure 33. AOI Fixation Duration Percentages in Segment 1

Figure 34. Mann-Whitney U-test Result for Seg-1 Fixation Duration Percentage

AOI	MW U-Statistic	p-value
Inst_Engine_P	20	0.009
Inst_Engine_S	17	0.054
Inst_Rudd_Ang	17	0.056
Own_St_P	15	0.066
ScrHdg_Lt	4	0.058
ScrHdg_Rt	4	0.058
Own_Bow_P	2.5	0.037

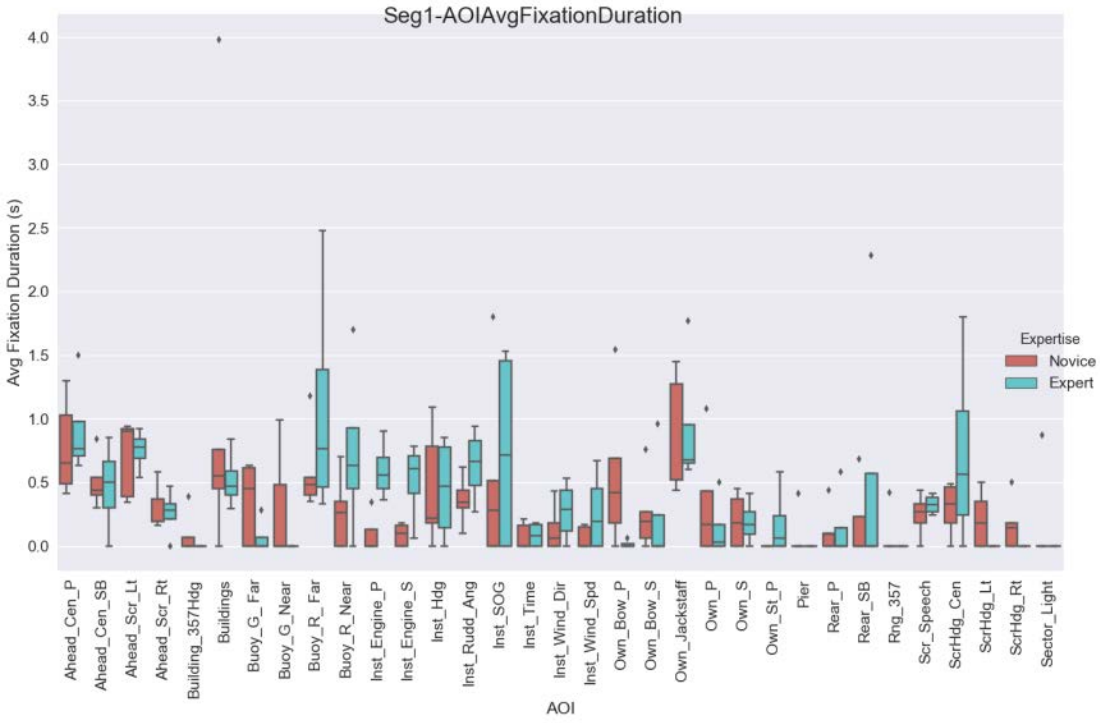


Figure 35. Average Fixation Duration in AOIs in Segment 1

Table 21. Mann-Whitney U-test Result for Seg-1 Avg Fixation Duration Percentage

AOI	MW U-Statistic	p-value
Inst_Engine_P	20	0.009
Inst_Engine_S	17	0.055
Own_St_P	15	0.066
ScrHdg_Lt	4	0.055
ScrHdg_Rt	4	0.055
Own_Bow_P	2.5	0.037

Segment 2 Metrics

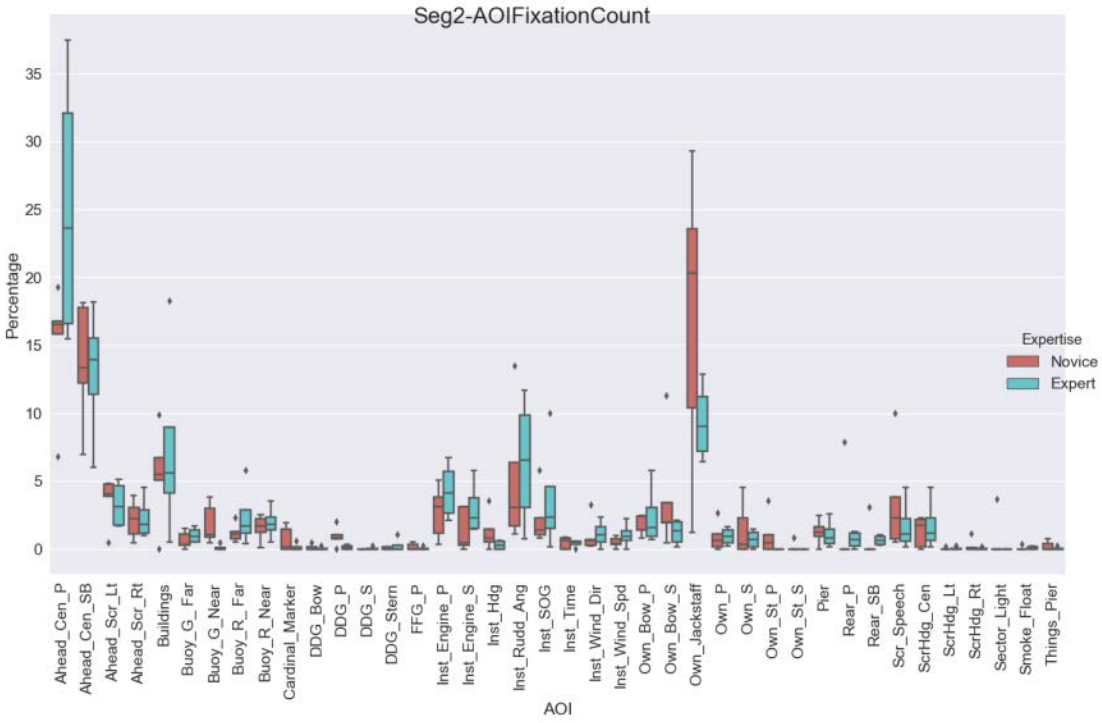


Figure 36. AOI Fixation Count Percentages in Segment 2

Table 22. Mann-Whitney U-test Result for Seg-2 Fixation Count Percentage

AOI	MW U-Statistic	p-value
Rear_SB	16	0.079
Buoy_G_Near	0	0.009
DDG_P	3	0.053
Own_St_P	4	0.055

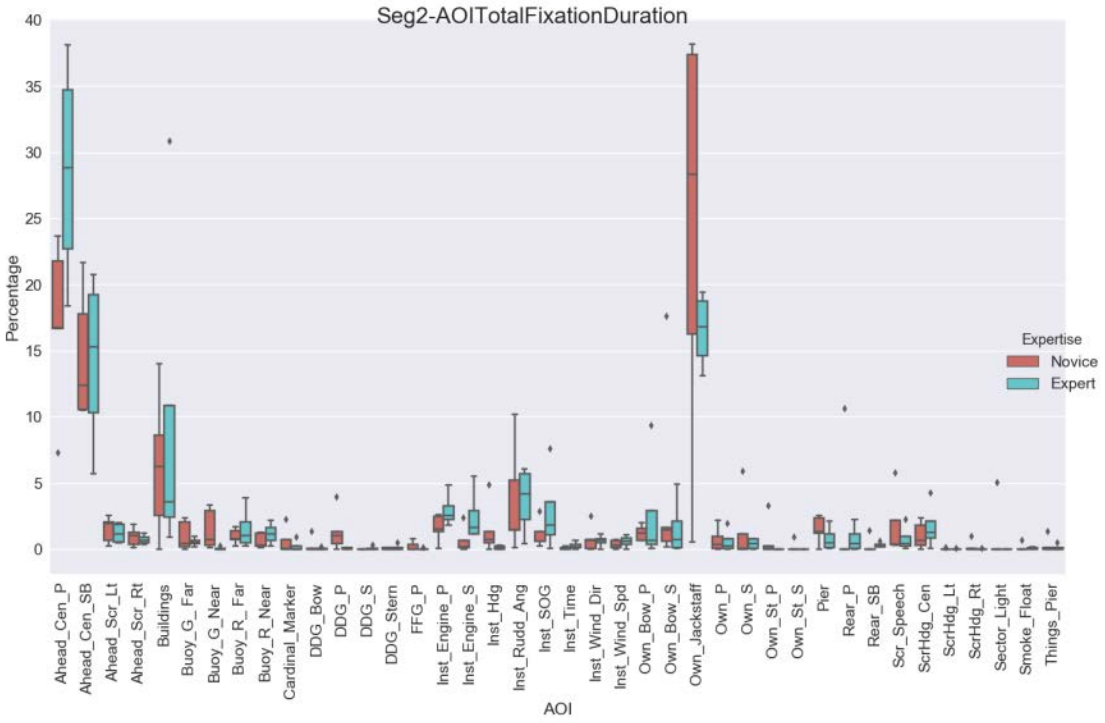


Figure 37. AOI Fixation Duration Percentages in Segment 2

Table 23. Mann-Whitney U-test Result for Seg-2 Fixation Duration Percentage

AOI	MW U-Statistic	p-value
Ahead_Cen_P	18	0.033
Inst_Engine_S	17	0.056
Inst_Engine_P	16	0.089
Rear_SB	16	0.079
Own_St_P	4	0.055
DDG_P	3	0.053
Inst_Hdg	3	0.053
Buoy_G_Near	1	0.017

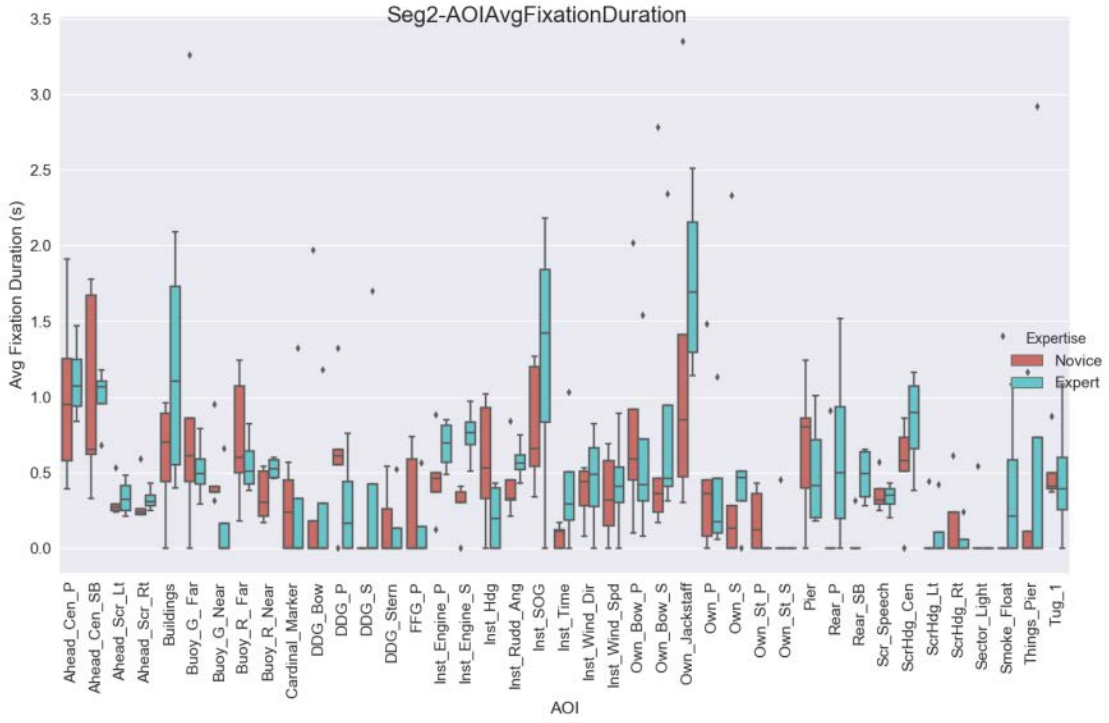


Figure 38. Average Fixation Duration in AOIs in Segment 2

Table 24. Mann-Whitney U-test Result for Seg-2 Average Fixation Duration Percentage

AOI	MW U-Statistic	p-value
Inst_Engine_S	20	0.009
Rear_SB	19	0.015
Buoy_R_Near	16	0.089
Inst_Time	16	0.085
Buoy_G_Near	4	0.084
Own_St_P	4	0.055

APPENDIX D. EYE SCAN AND AOI TIME DISTRIBUTION GRAPHS

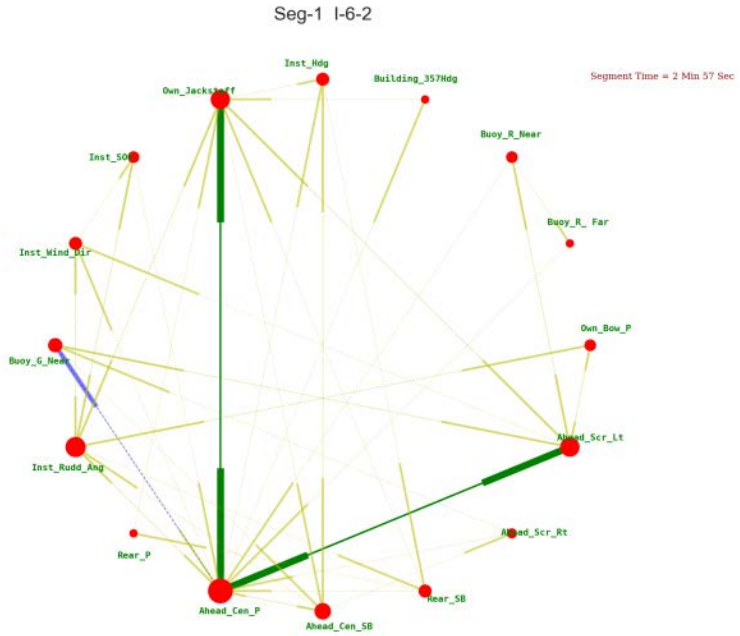


Figure 39. I-6-2 Eye Scan Network Graph in Segment 1

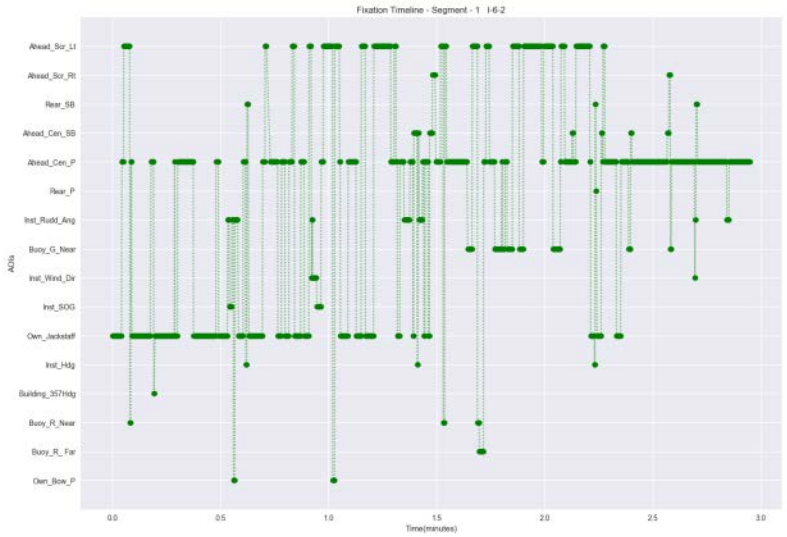


Figure 40. I-6-2 Eye Scan Timeline Graph in Segment 1

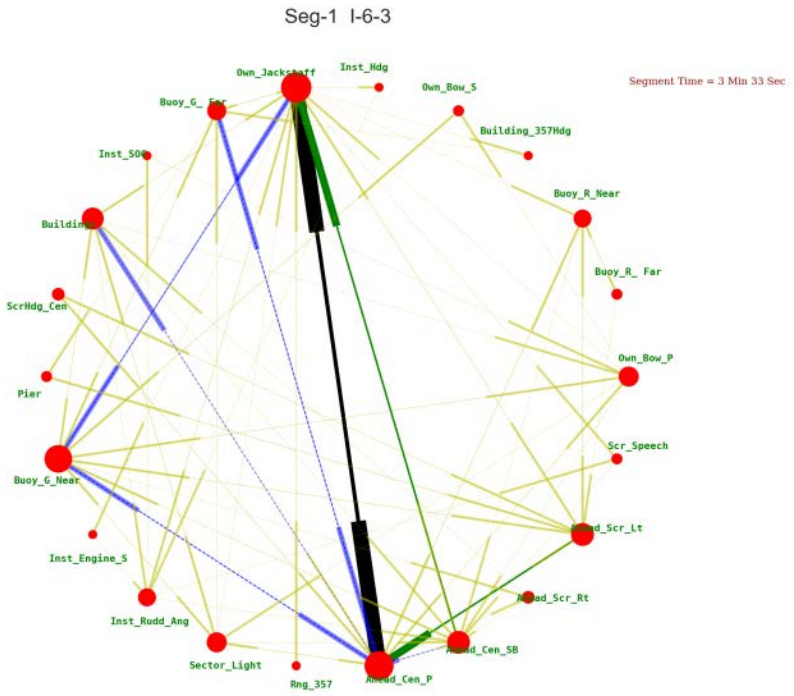


Figure 41. I-6-3 Eye Scan Network Graph in Segment 1

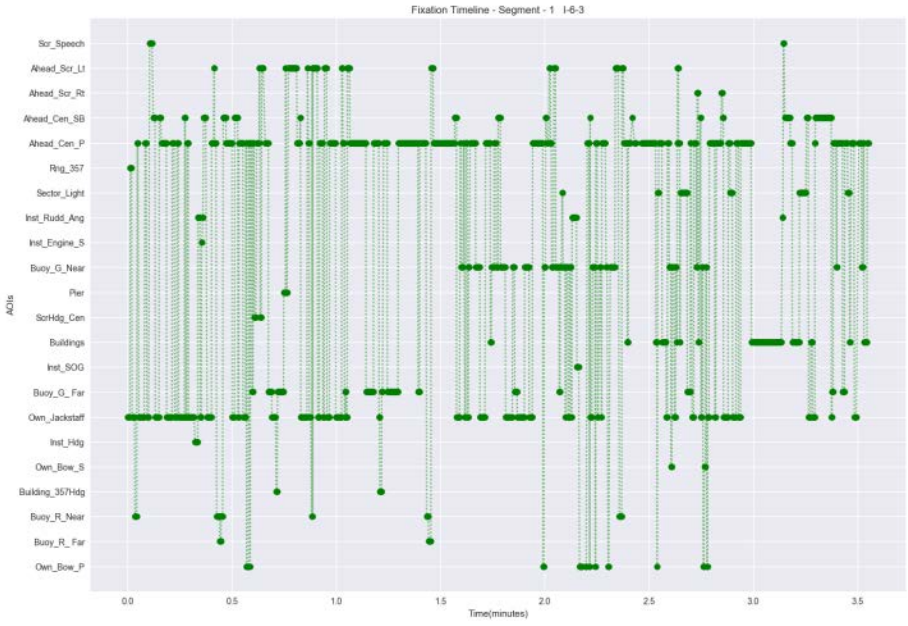


Figure 42. I-6-3 Eye Scan Timeline Graph in Segment 1

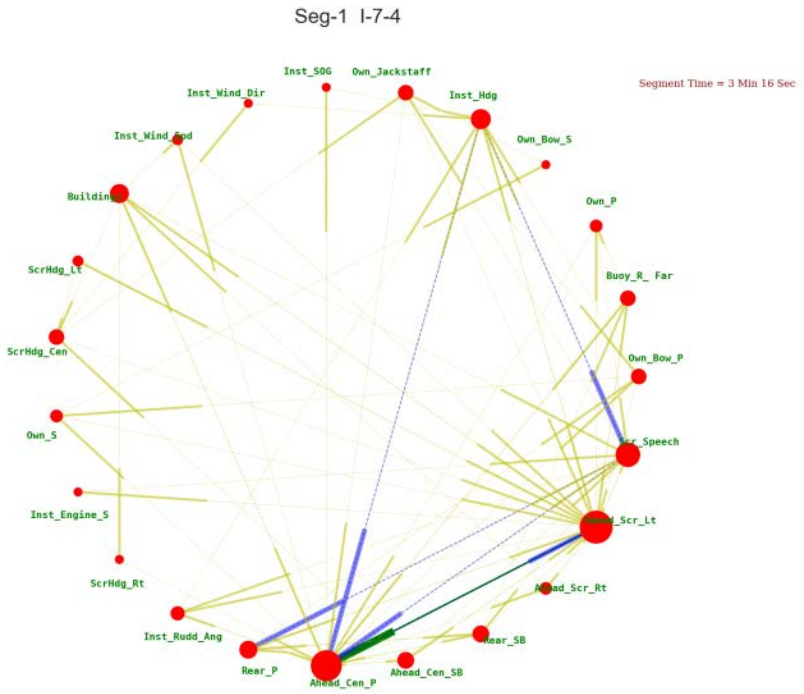


Figure 43. I-7-4 Eye Scan Network Graph in Segment 1

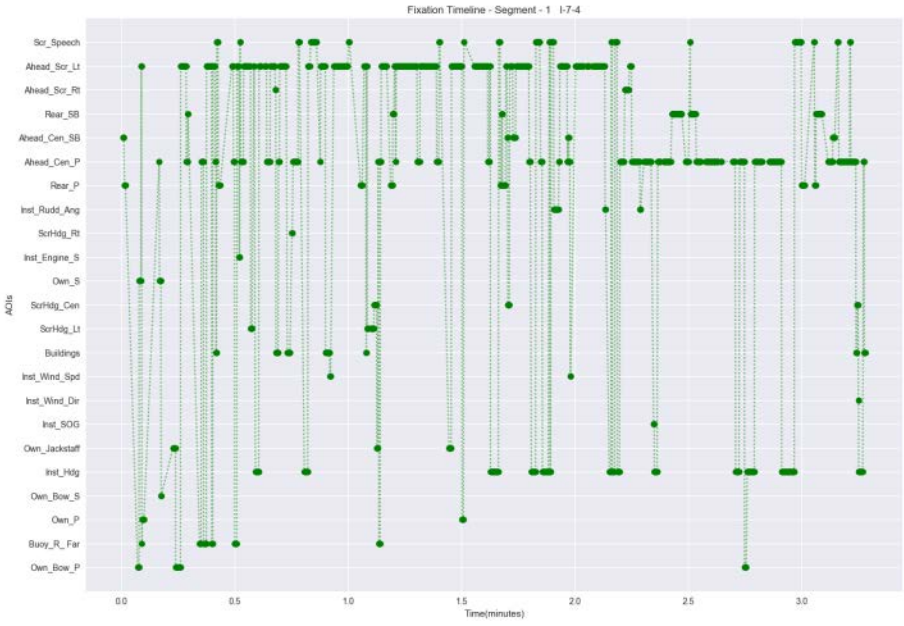


Figure 44. I-7-4 Eye Scan Timeline Graph in Segment 1

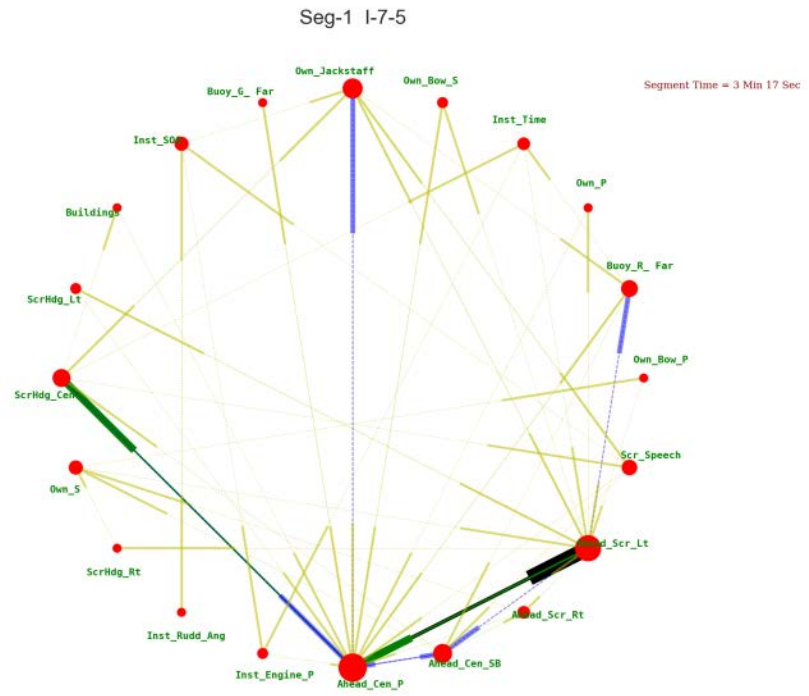


Figure 45. I-7-5 Eye Scan Network Graph in Segment 1

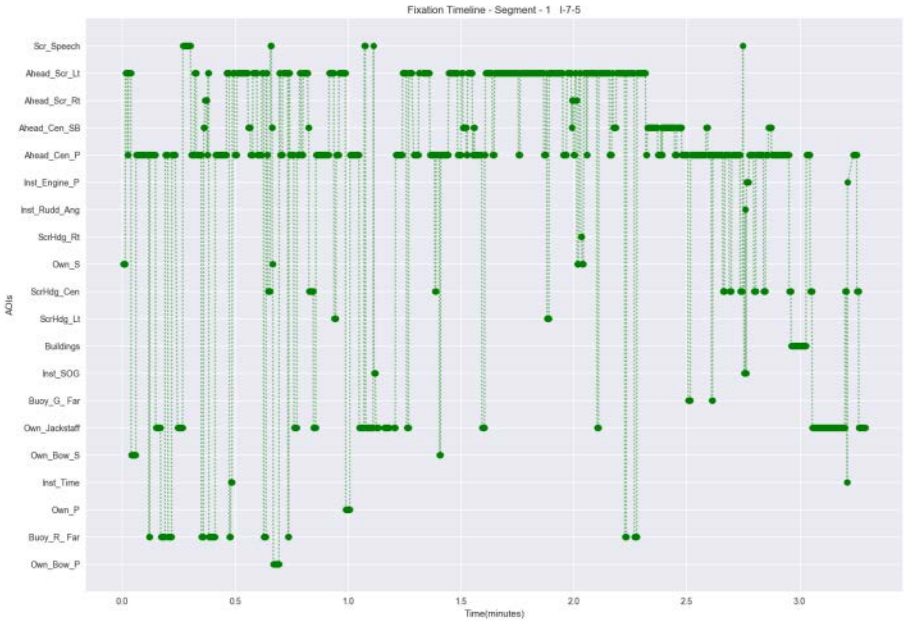


Figure 46. I-7-5 Eye Scan Timeline Graph in Segment 1

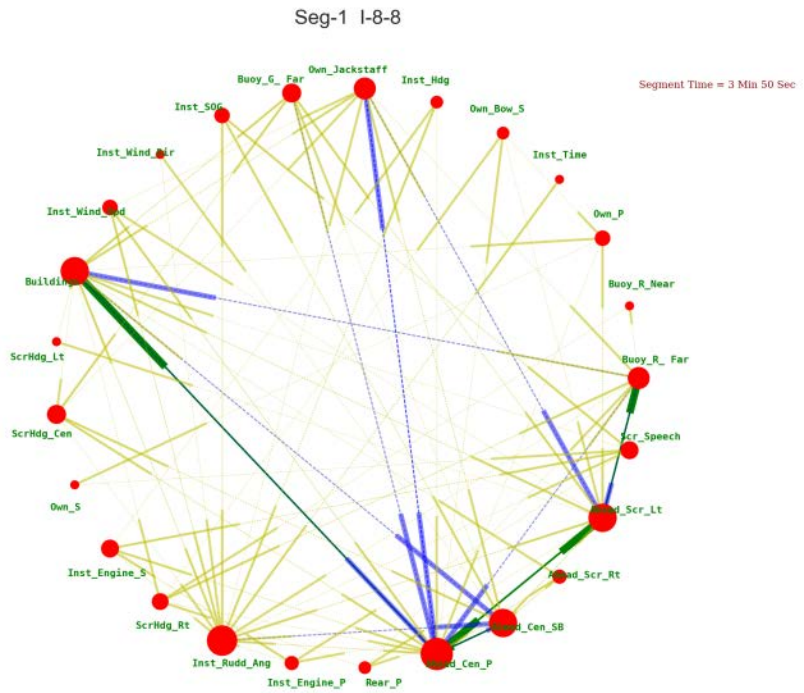


Figure 47. I-8-8 Eye Scan Network Graph in Segment 1

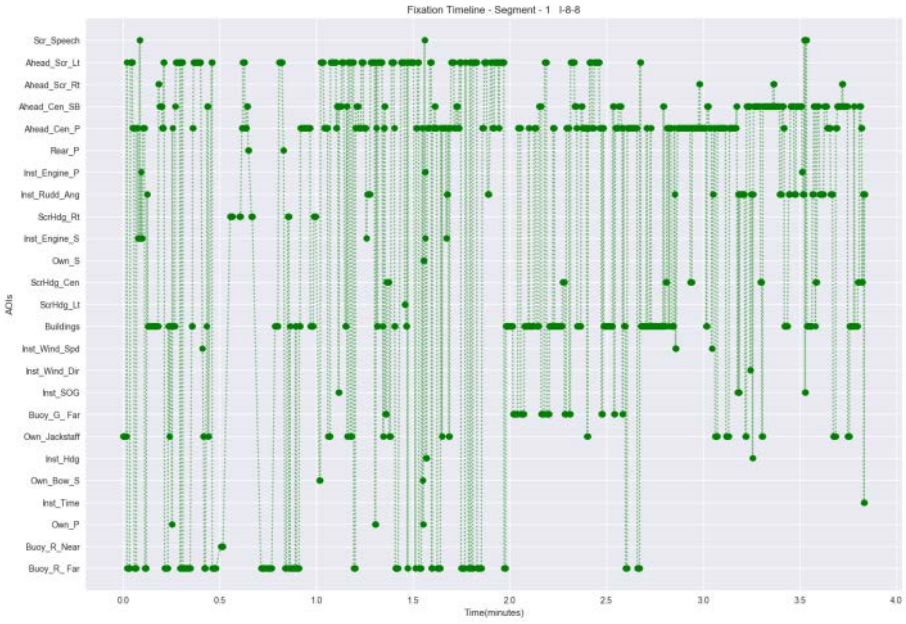


Figure 48. I-8-8 Eye Scan Timeline Graph in Segment 1

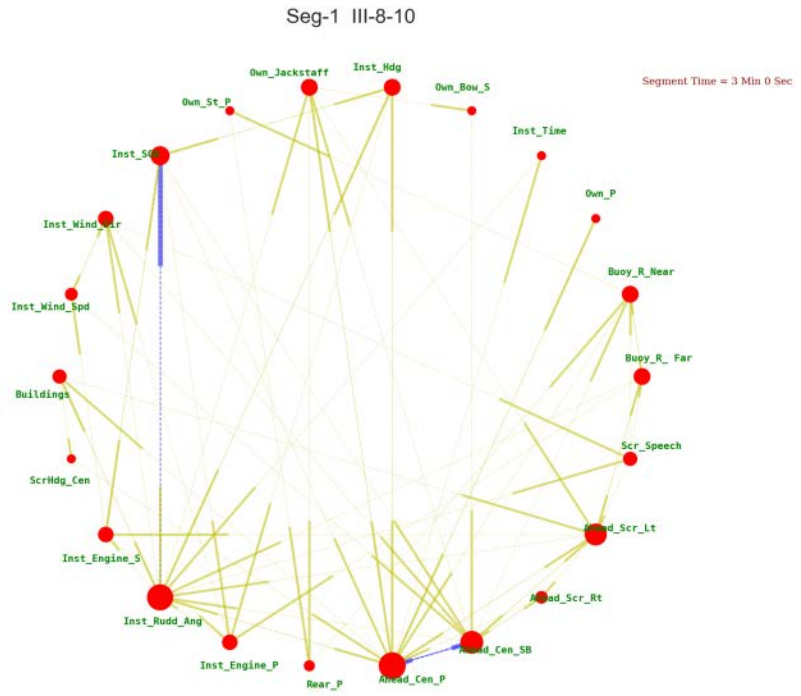


Figure 49. III-8-10 Eye Scan Network Graph in Segment 1

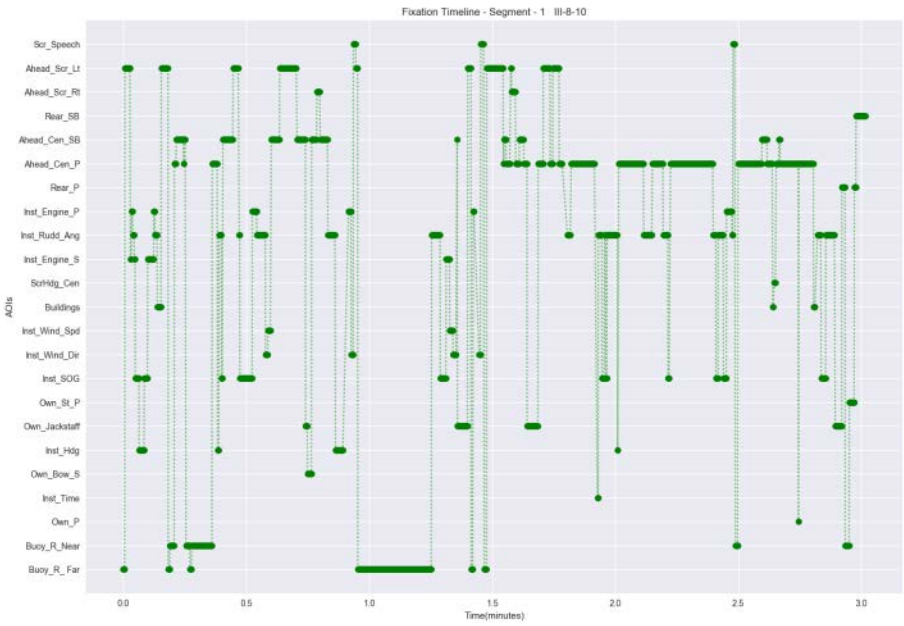


Figure 50. III-8-10 Eye Scan Timeline Graph in Segment 1

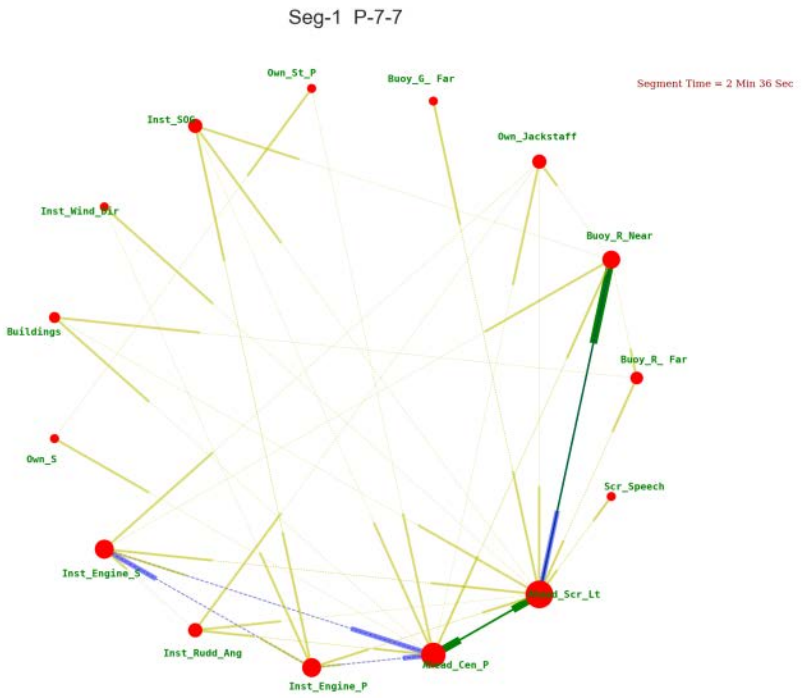


Figure 51. P-7-7 Eye Scan Network Graph in Segment 1

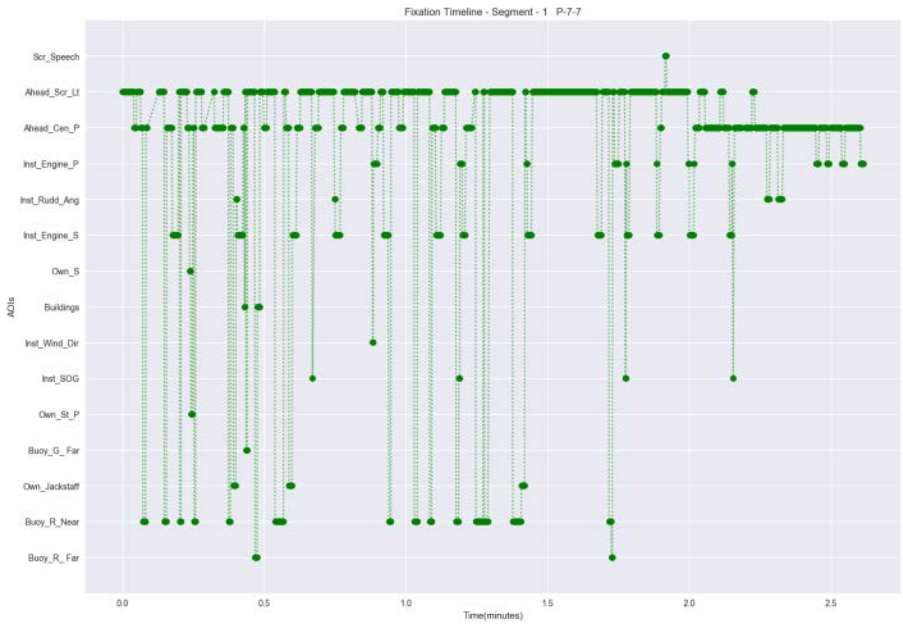


Figure 52. P-7-7 Eye Scan Timeline Graph in Segment 1

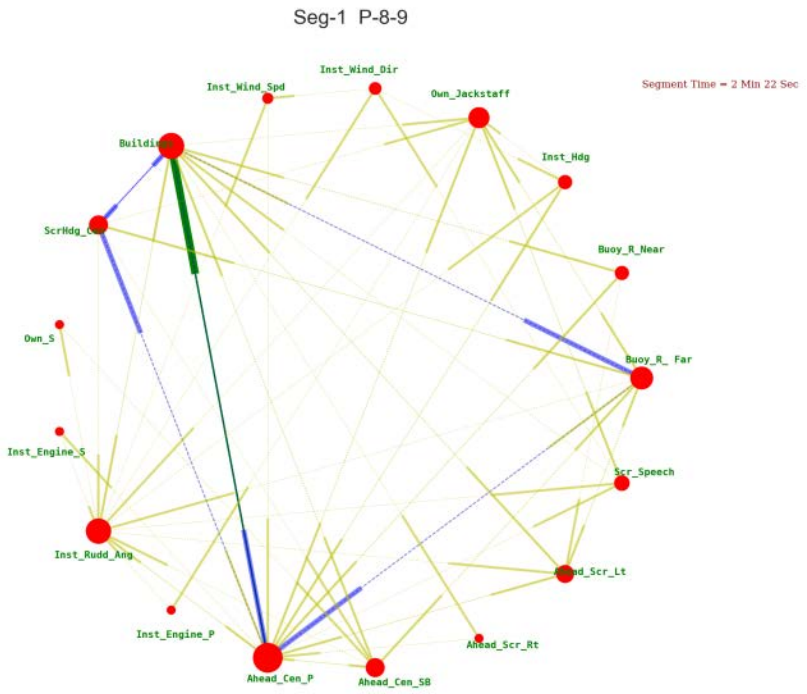


Figure 53. P-8-9 Eye Scan Network Graph in Segment 1

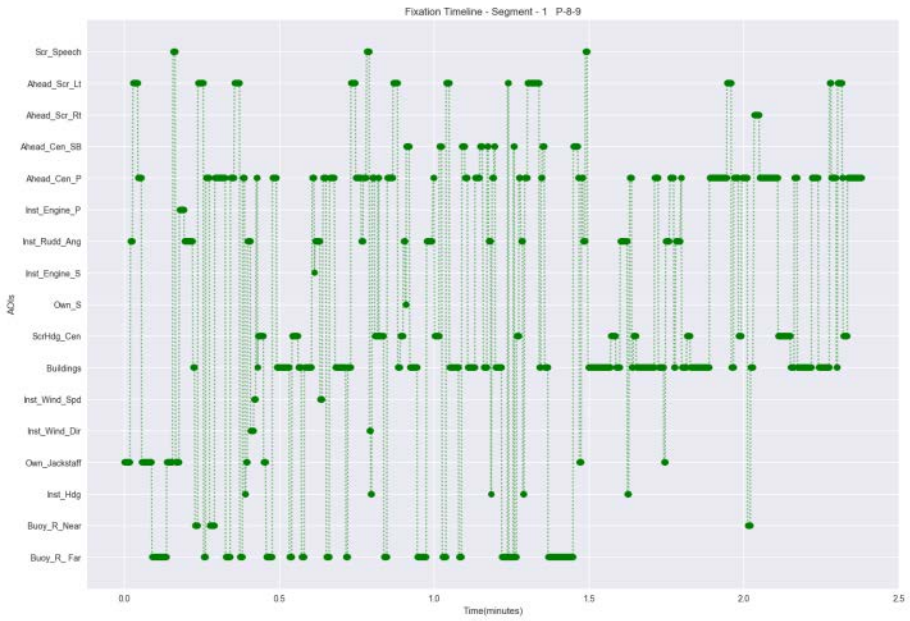


Figure 54. P-8-9 Eye Scan Timeline Graph in Segment 1

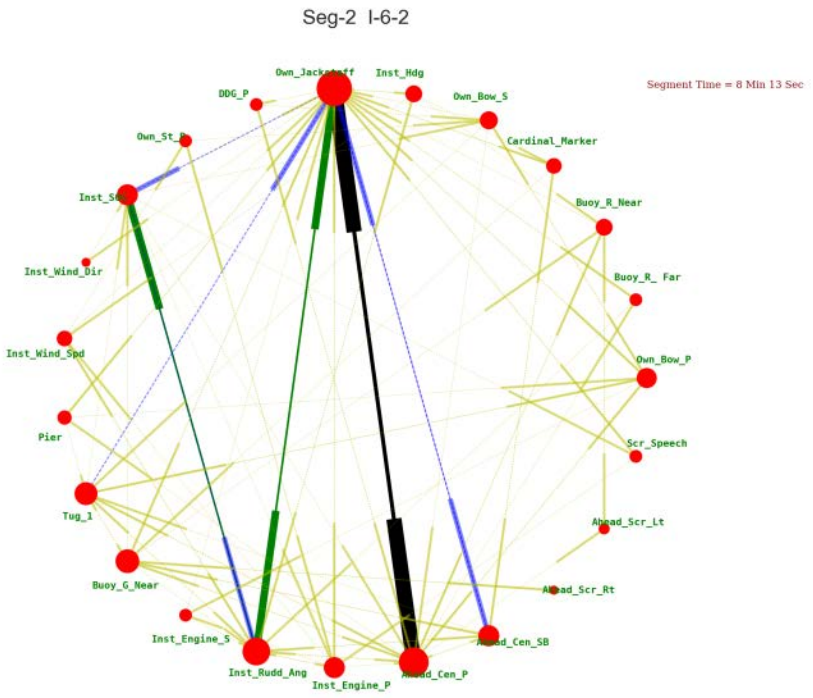


Figure 55. I-6-2 Eye Scan Network Graph in Segment 2

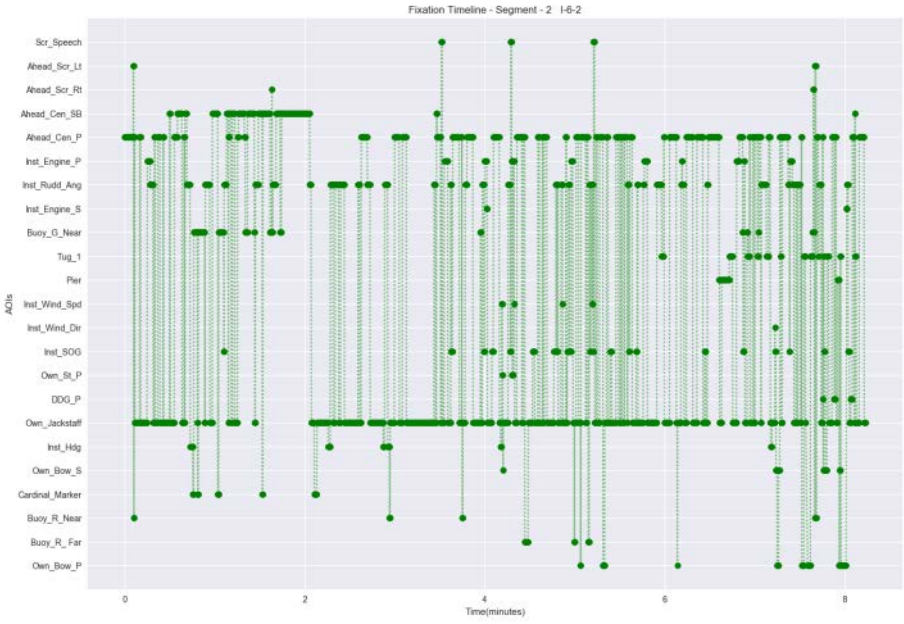


Figure 56. I-6-2 Eye Scan Timeline Graph in Segment 2

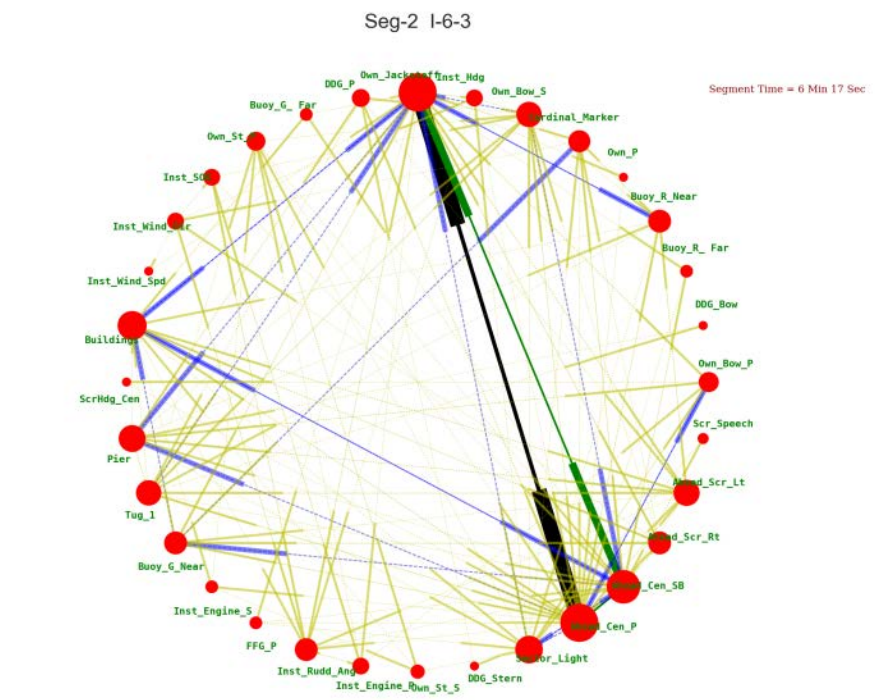


Figure 57. I-6-3 Eye Scan Network Graph in Segment 2

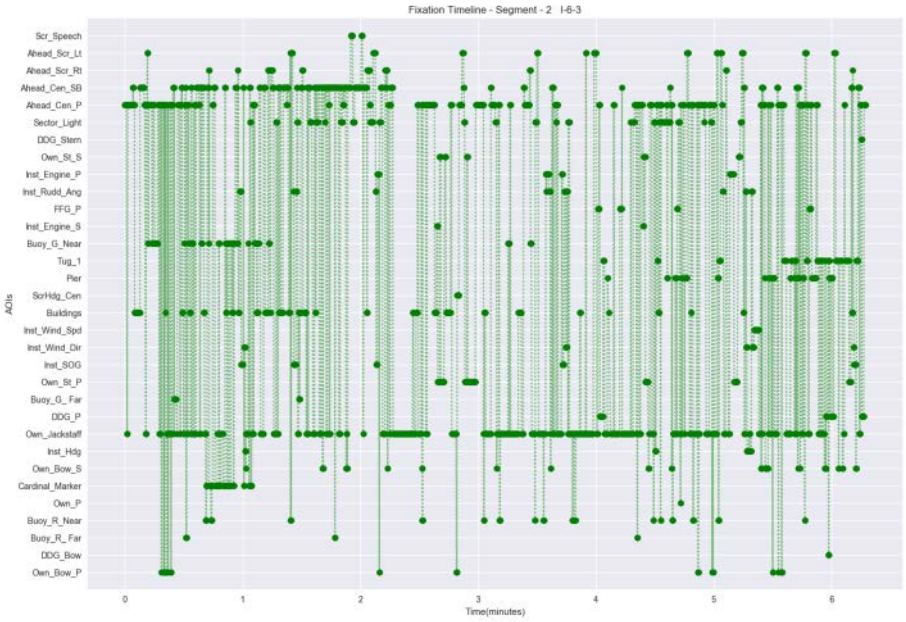


Figure 58. I-6-3 Eye Scan Timeline Graph in Segment 2

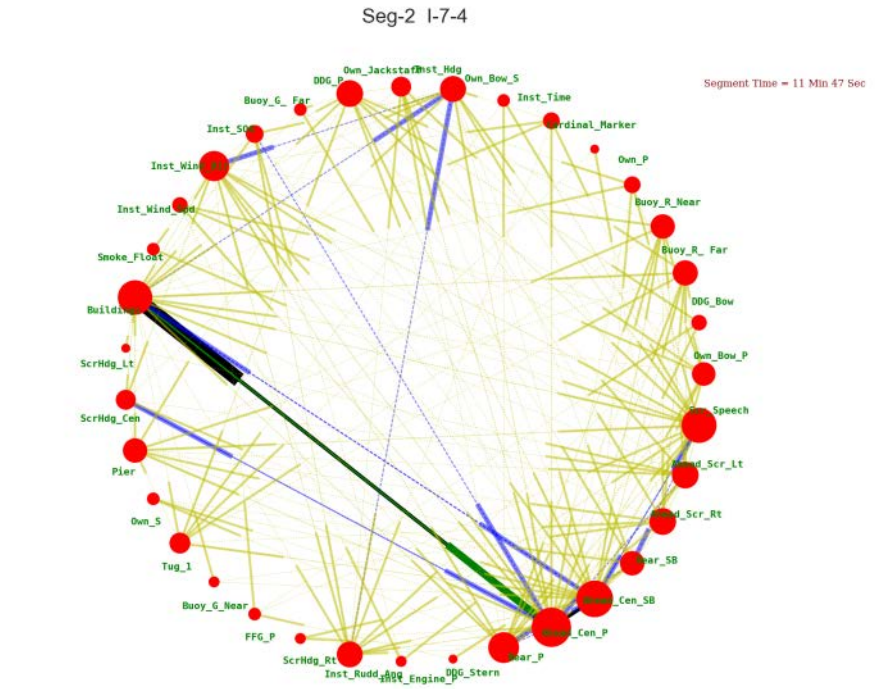


Figure 59. I-7-4 Eye Scan Network Graph in Segment 2

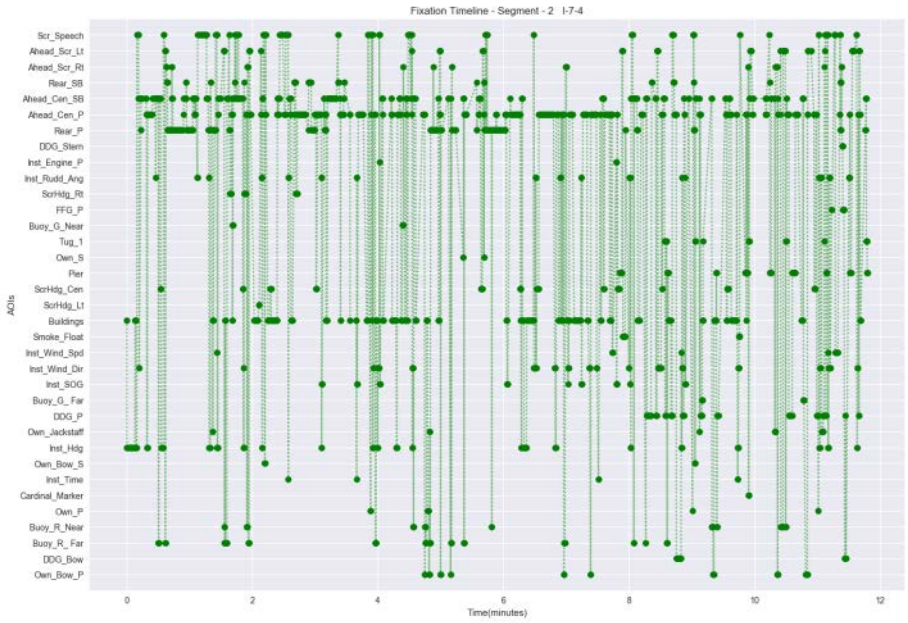


Figure 60. I-7-4 Eye Scan Timeline Graph in Segment 2

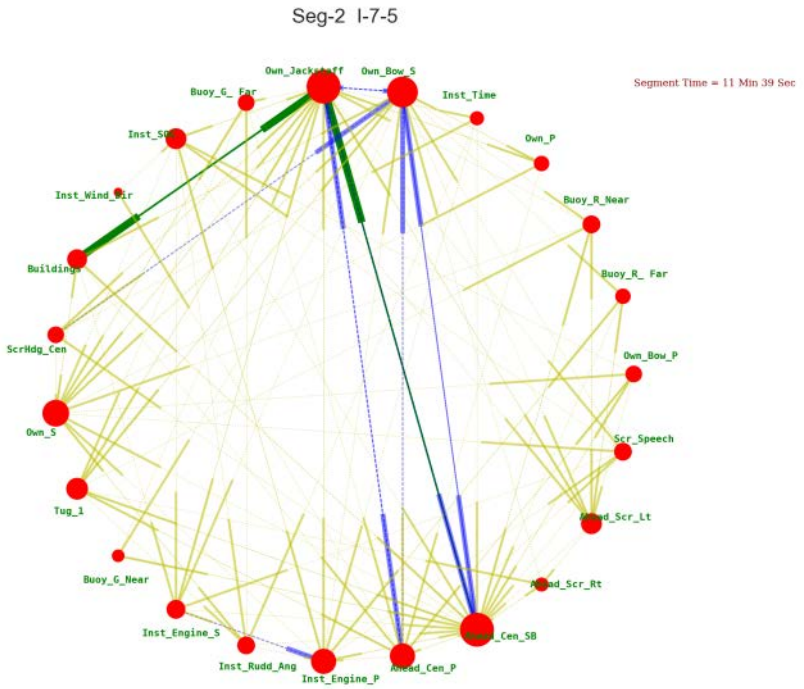


Figure 61. I-7-5 Eye Scan Network Graph in Segment 2

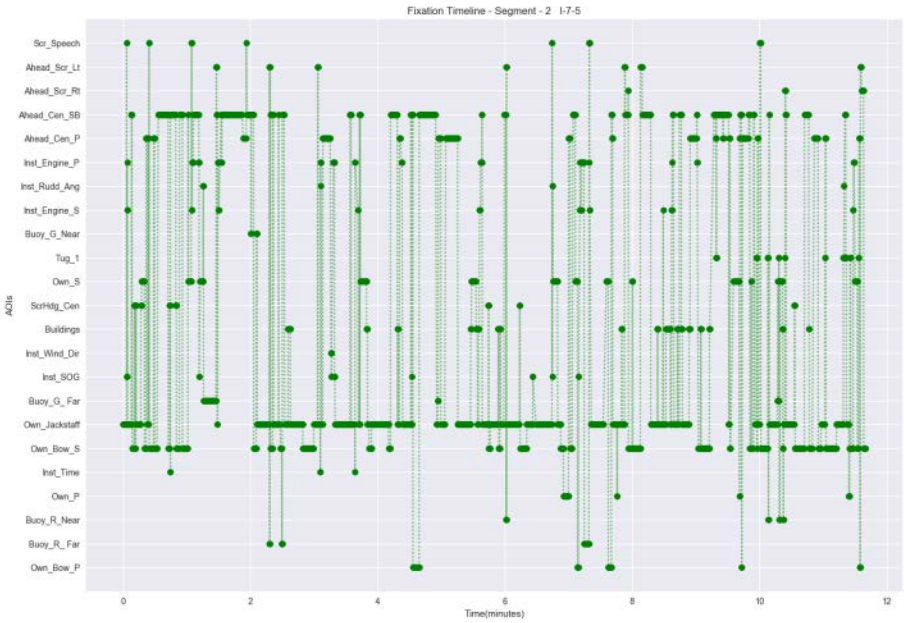


Figure 62. I-7-5 Eye Scan Timeline Graph in Segment 2

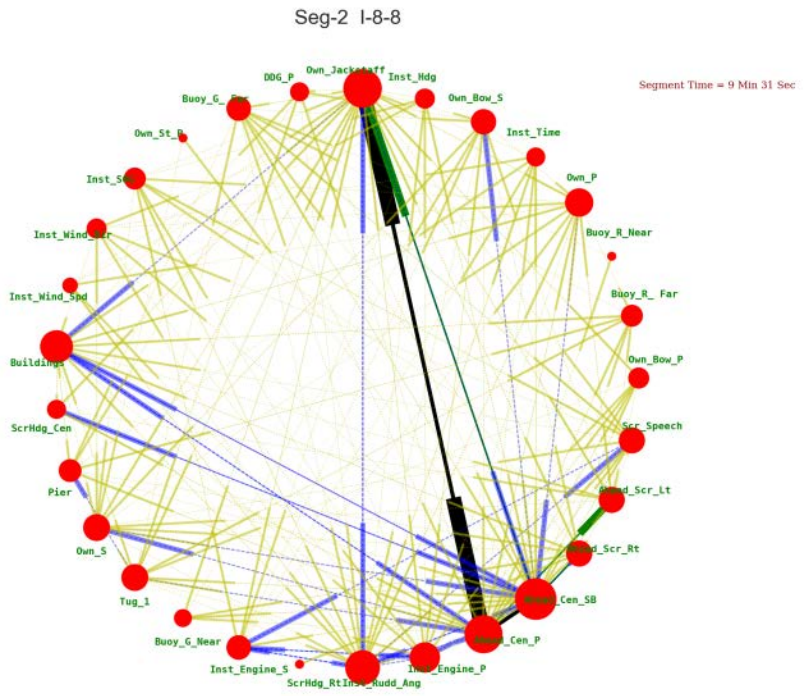


Figure 63. I-8-8 Eye Scan Network Graph in Segment 2

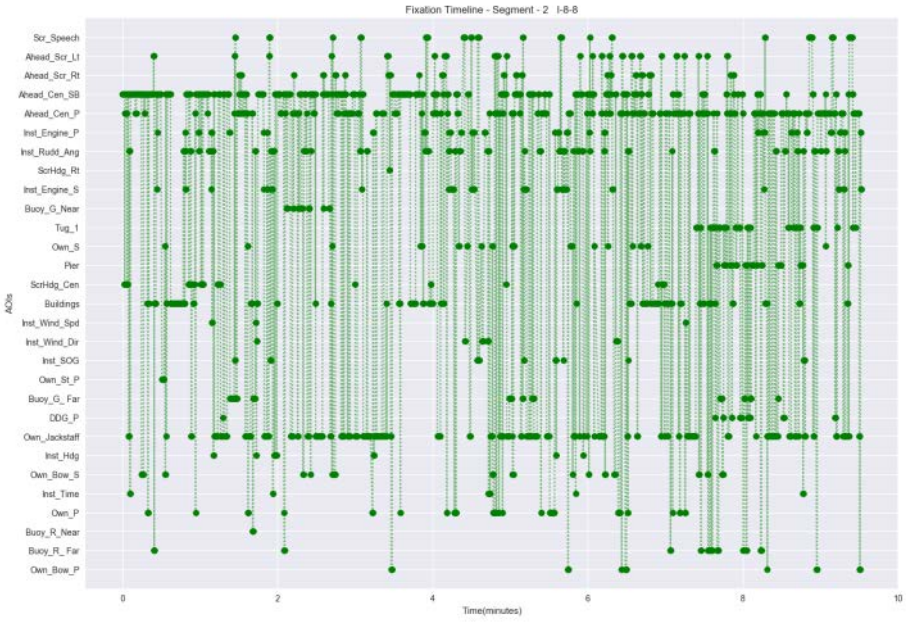


Figure 64. I-8-8 Eye Scan Timeline Graph in Segment 2

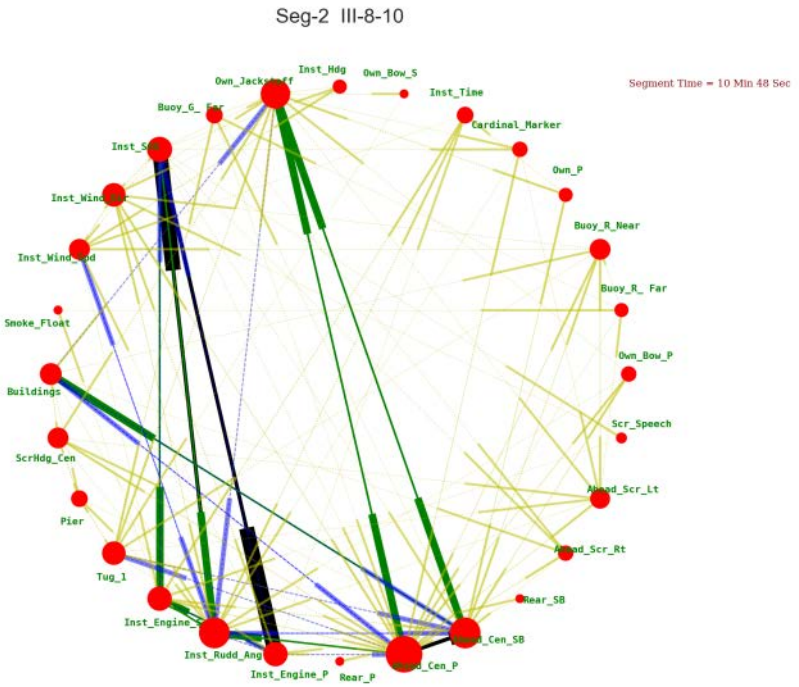


Figure 65. III-8-10 Eye Scan Network Graph in Segment 2

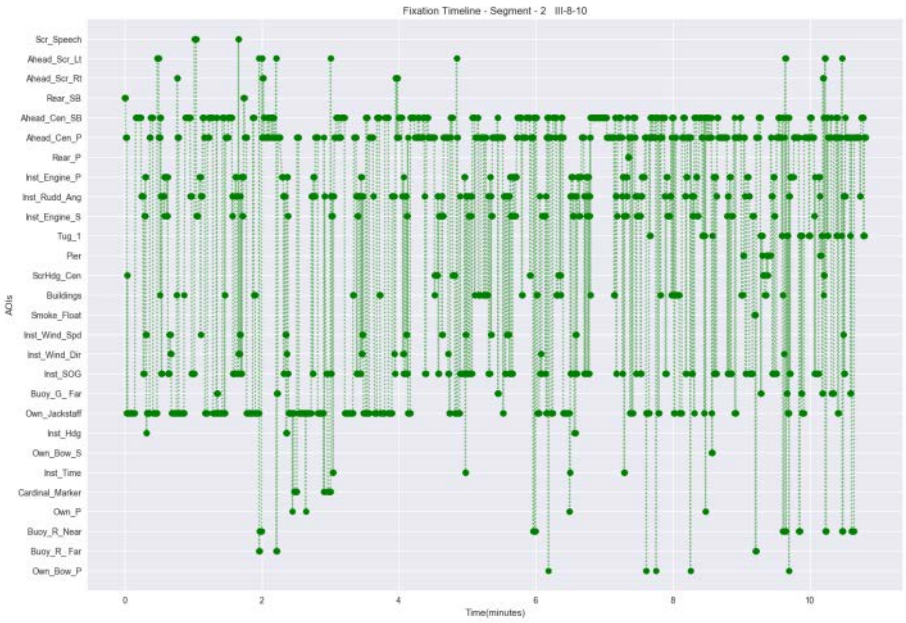


Figure 66. III-8-10 Eye Scan Timeline Graph in Segment 2

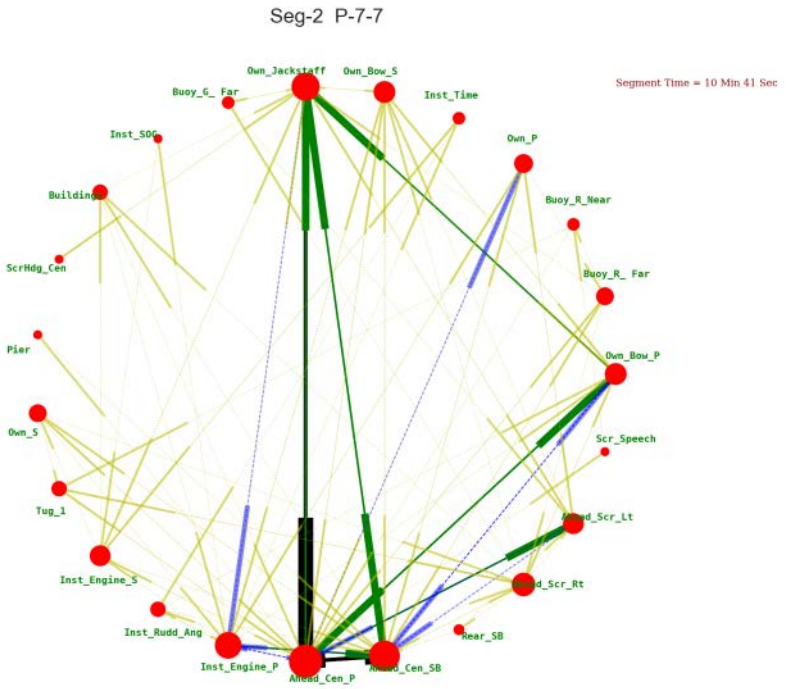


Figure 67. P-7-7 Eye Scan Network Graph in Segment 2

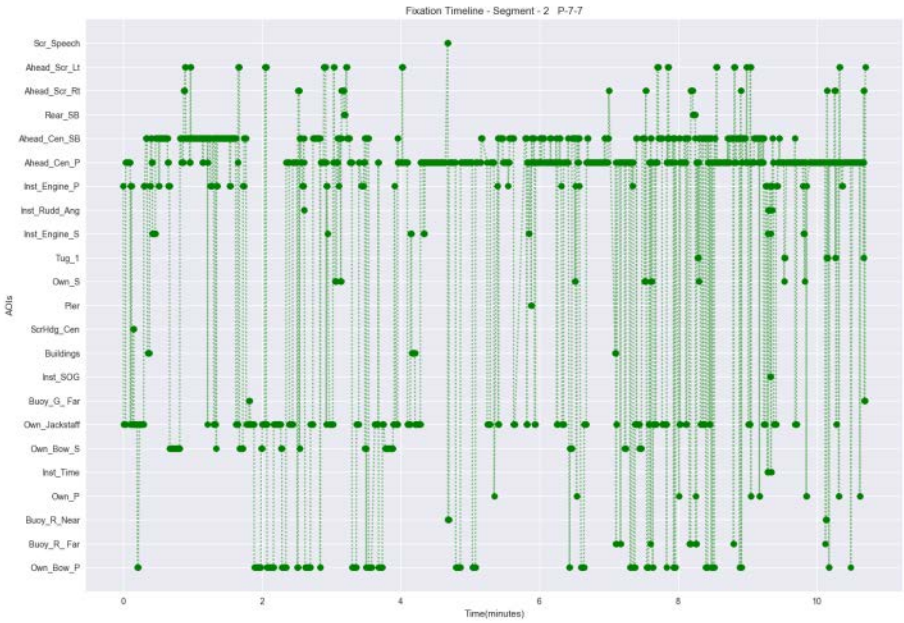


Figure 68. P-7-7 Eye Scan Timeline Graph in Segment 2

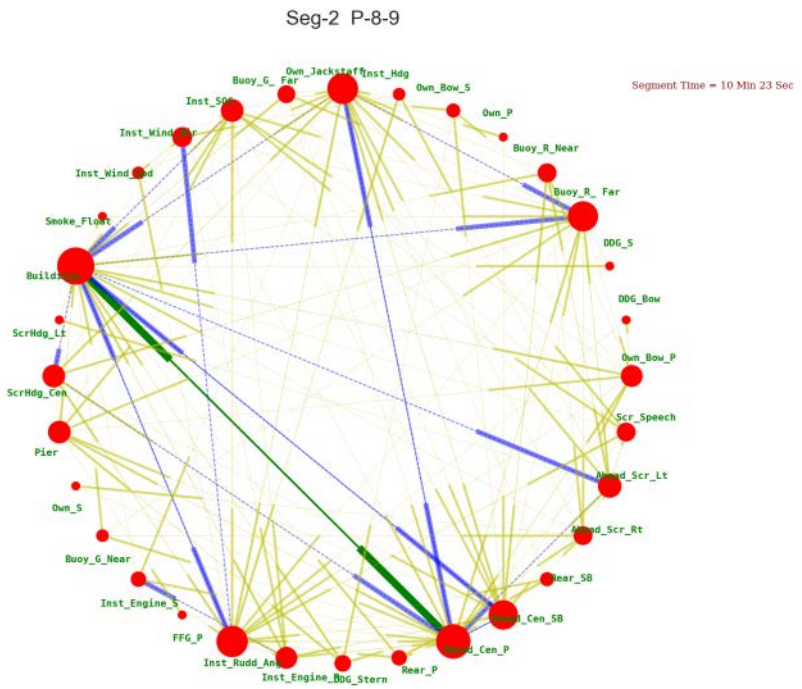


Figure 69. P-8-9 Eye Scan Network Graph in Segment 2

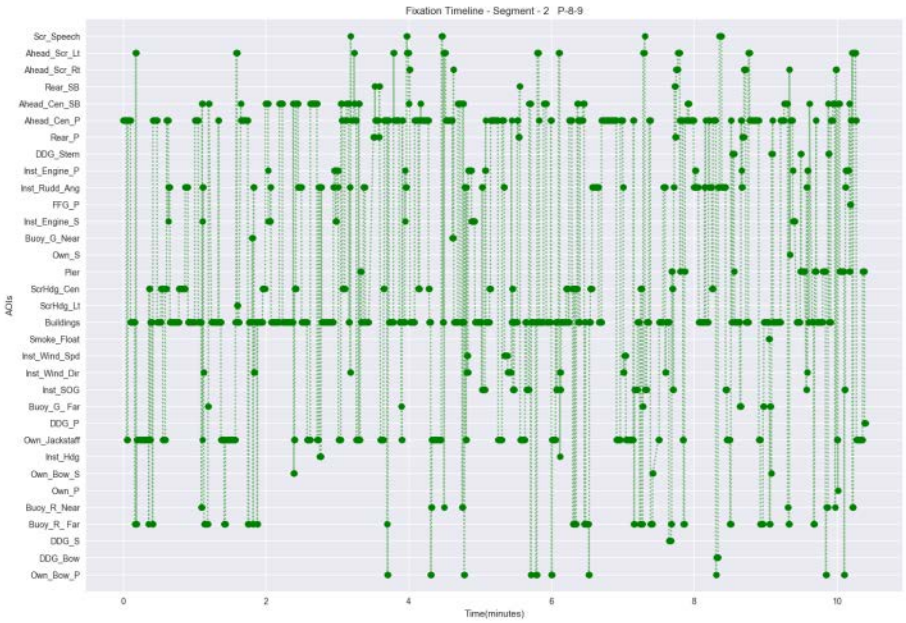


Figure 70. P-8-9 Eye Scan Timeline Graph in Segment 2

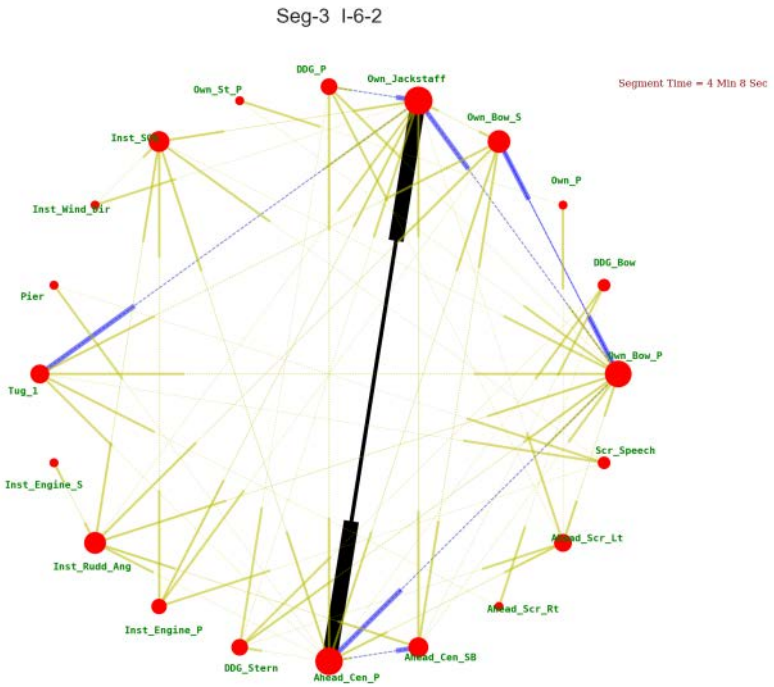


Figure 71. I-6-2 Eye Scan Network Graph in Segment 3

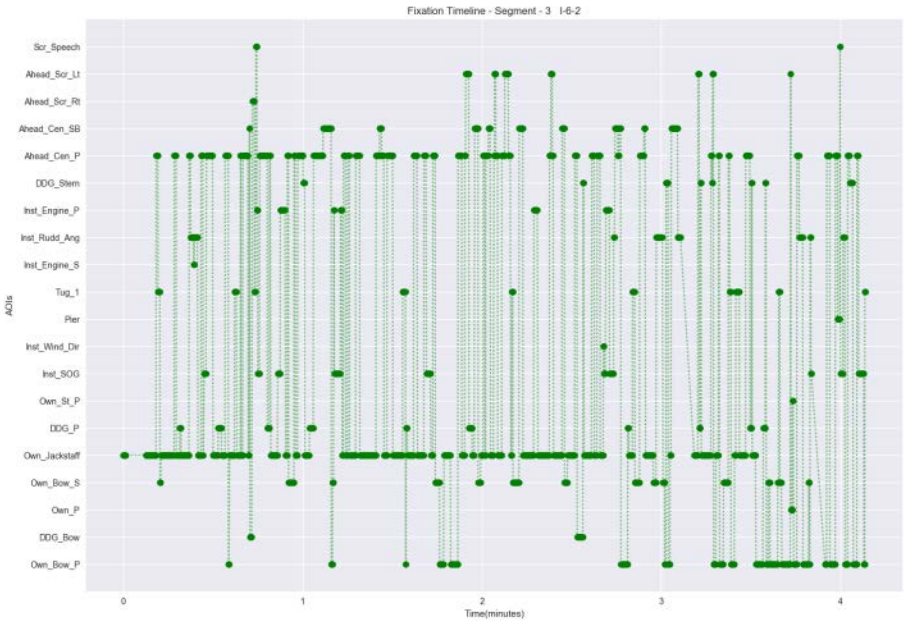


Figure 72. I-6-2 Eye Scan Timeline Graph in Segment 3

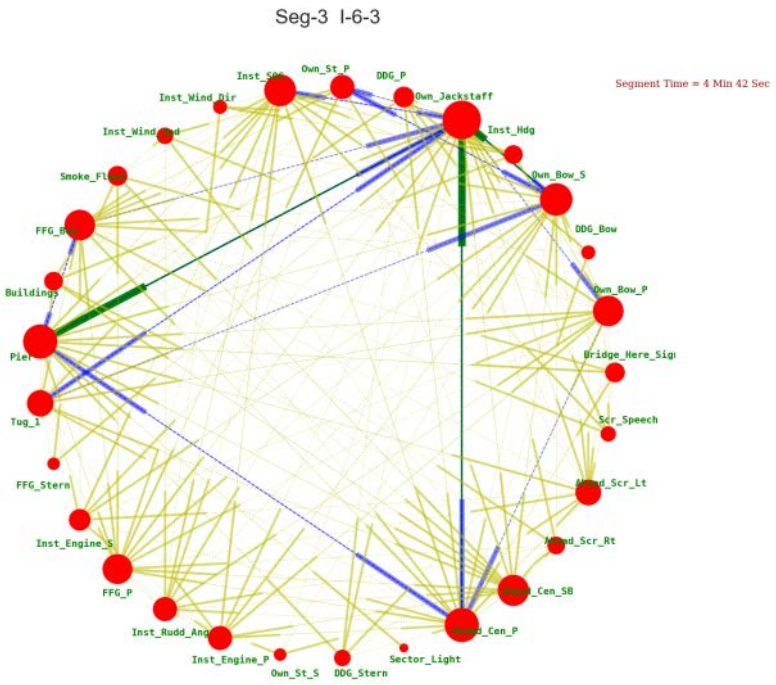


Figure 73. I-6-3 Eye Scan Network Graph in Segment 3

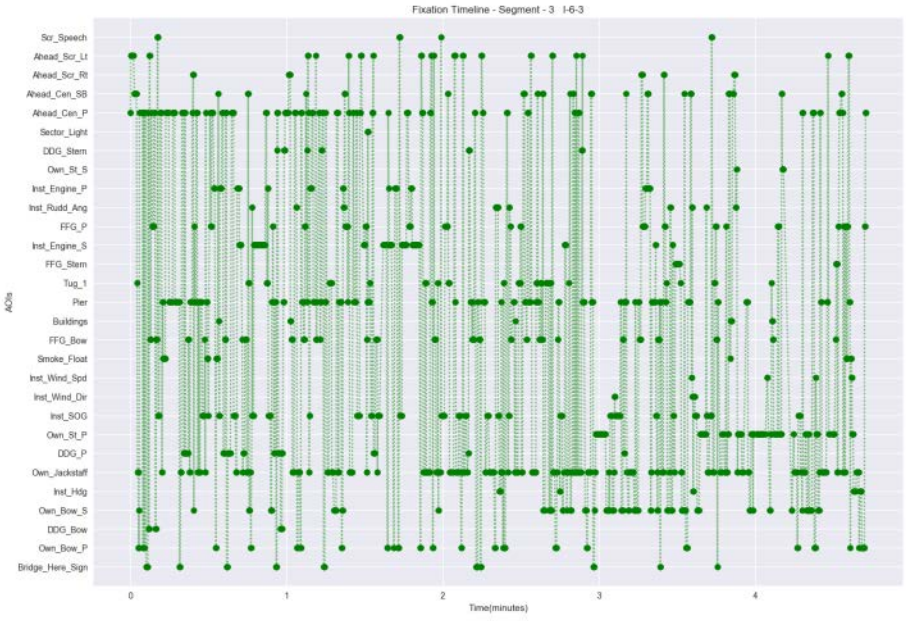


Figure 74. I-6-3 Eye Scan Timeline Graph in Segment 3

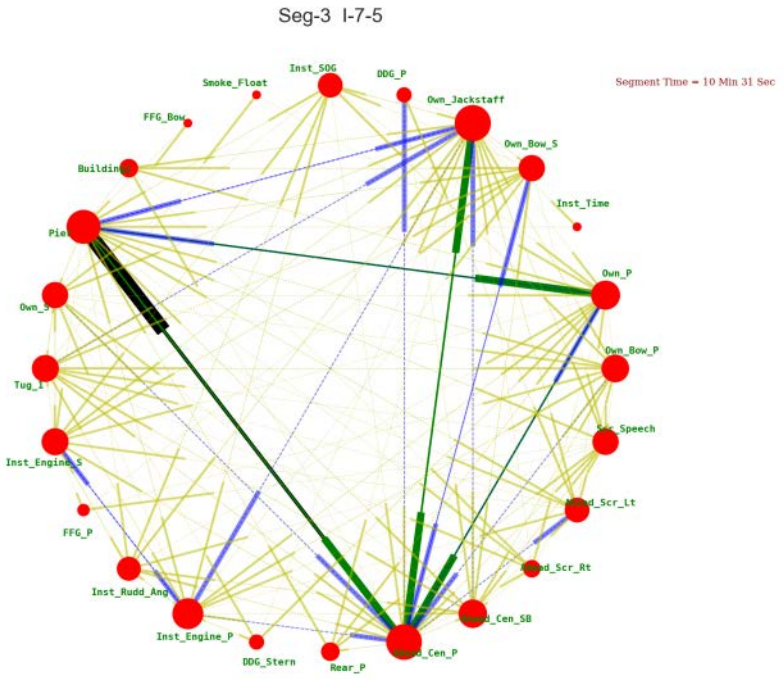


Figure 75. I-7-5 Eye Scan Network Graph in Segment 3

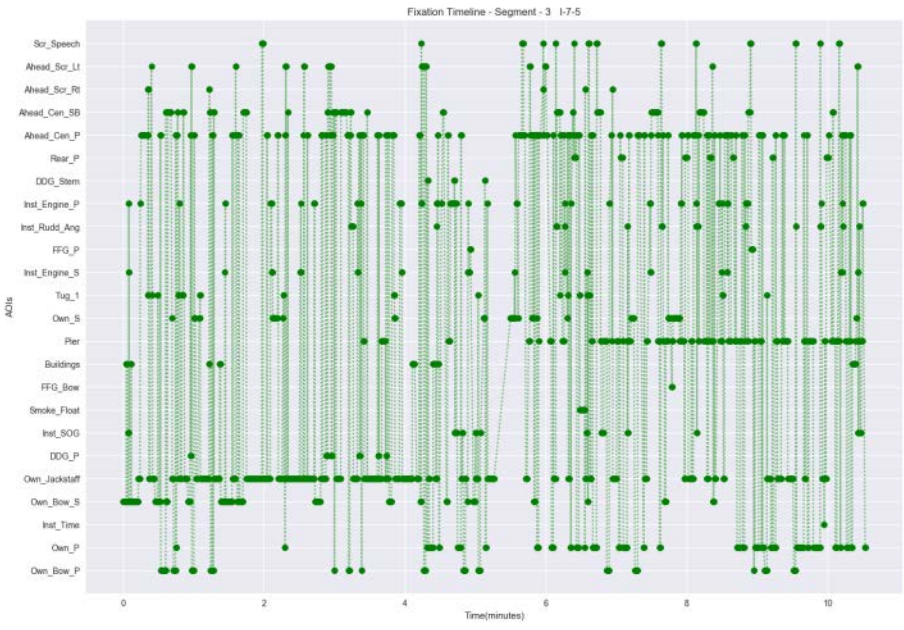


Figure 76. I-7-5 Eye Scan Timeline Graph in Segment 3

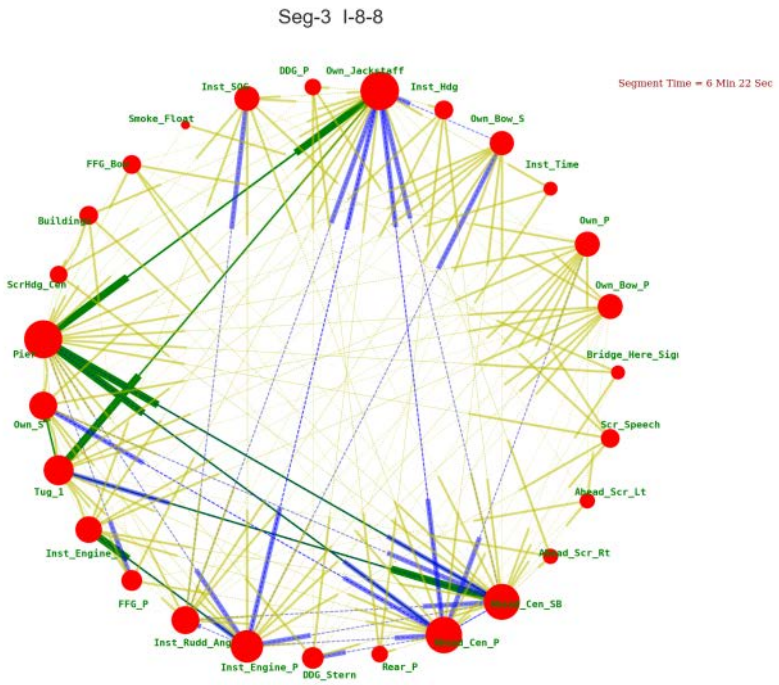


Figure 77. I-8-8 Eye Scan Network Graph in Segment 3

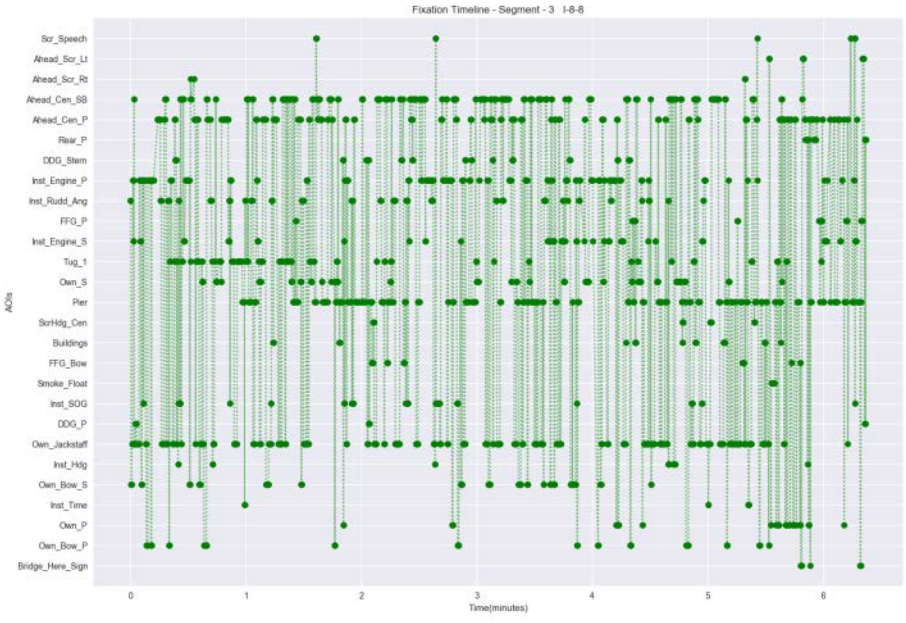


Figure 78. I-8-8 Eye Scan Timeline Graph in Segment 3

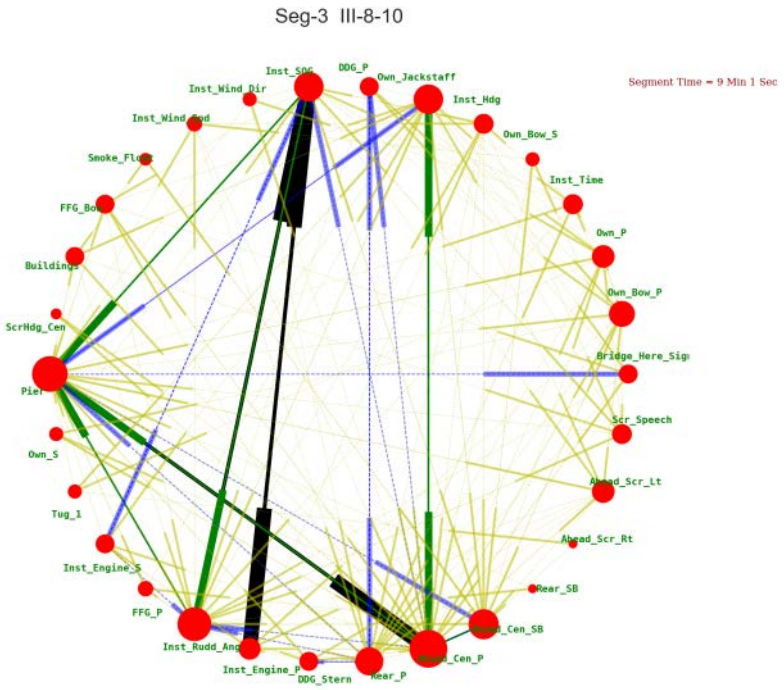


Figure 79. III-8-10 Eye Scan Network Graph in Segment 3

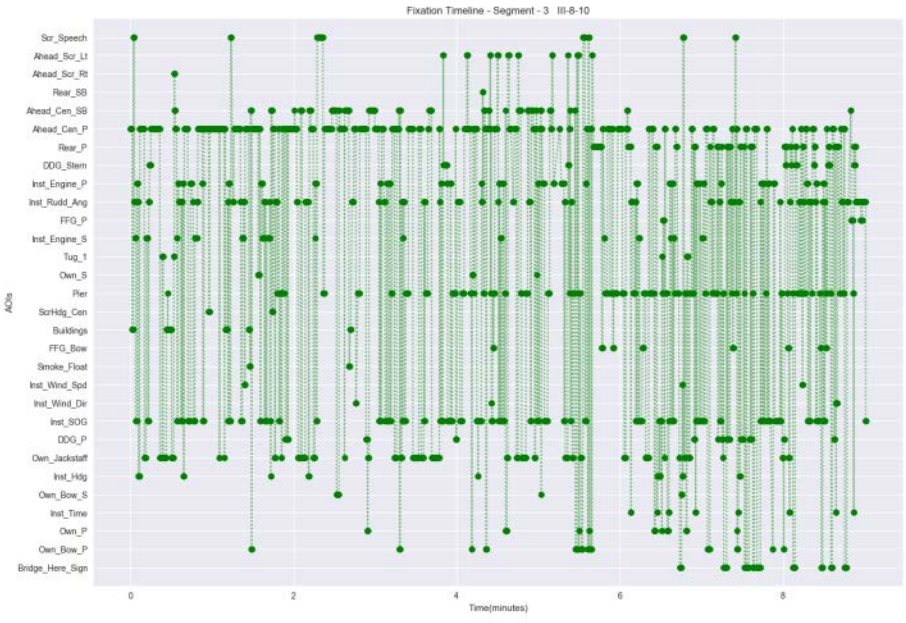


Figure 80. III-8-10 Eye Scan Timeline Graph in Segment 3

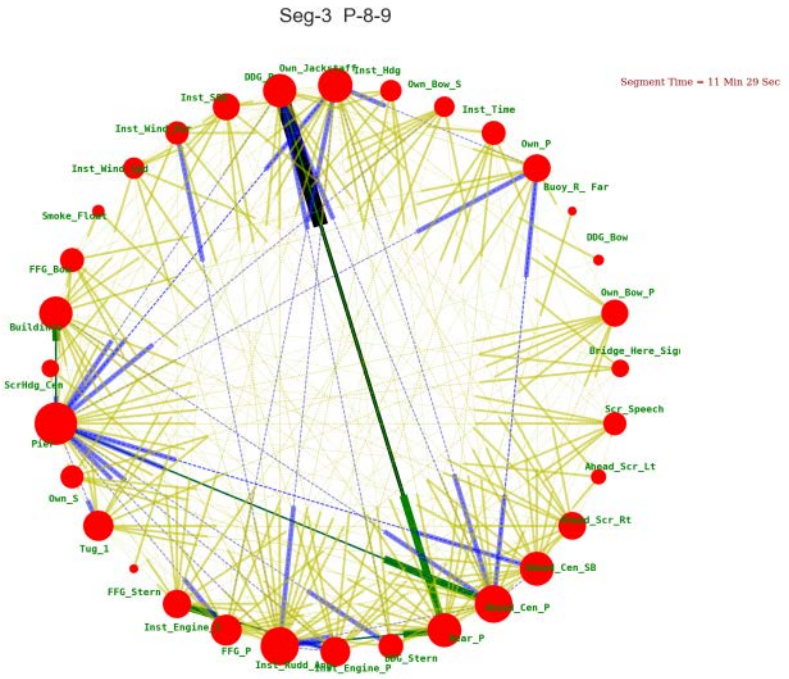


Figure 81. P-8-9 Eye Scan Network Graph in Segment 3

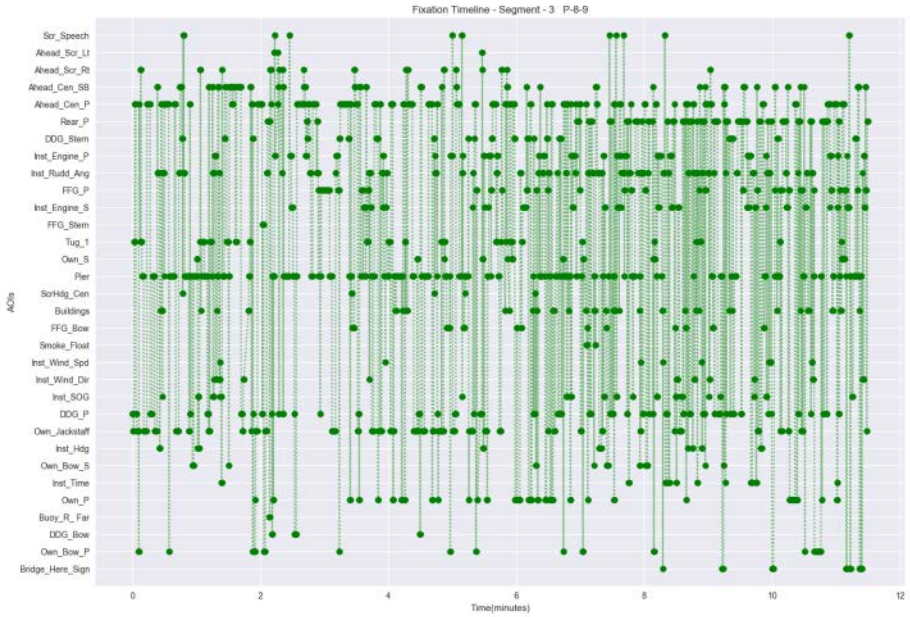


Figure 82. P-8-9 Eye Scan Timeline Graph in Segment 3

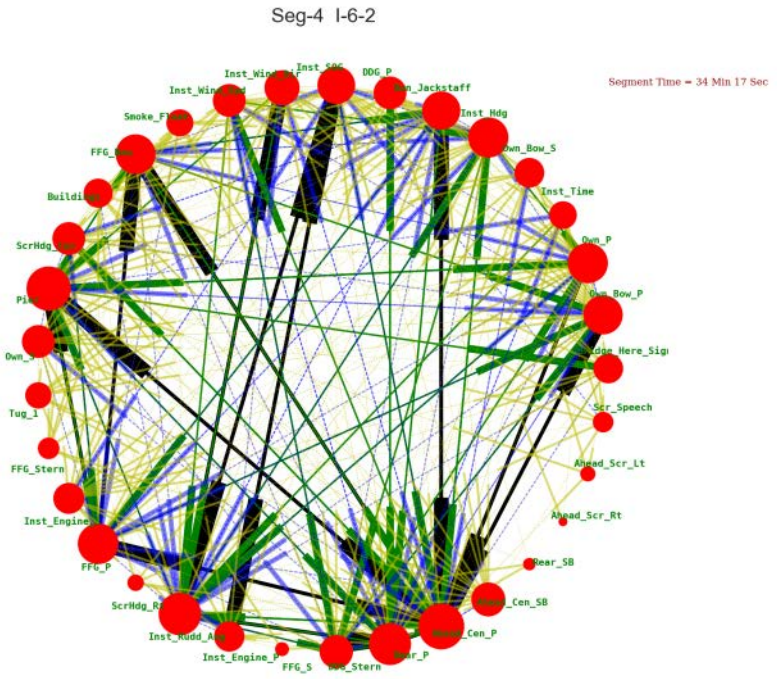


Figure 83. I-6-2 Eye Scan Network Graph in Segment 4

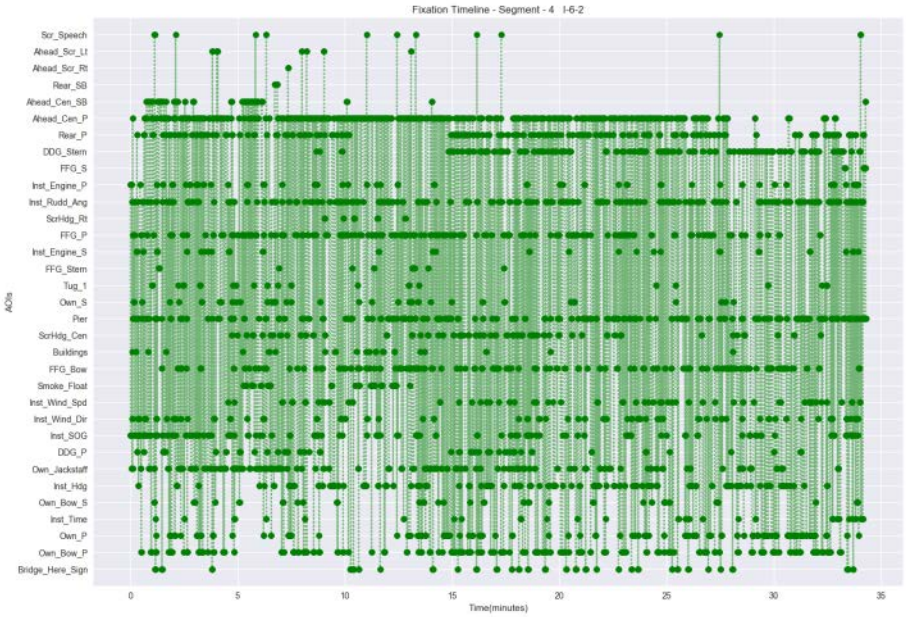


Figure 84. I-6-2 Eye Scan Timeline Graph in Segment 4

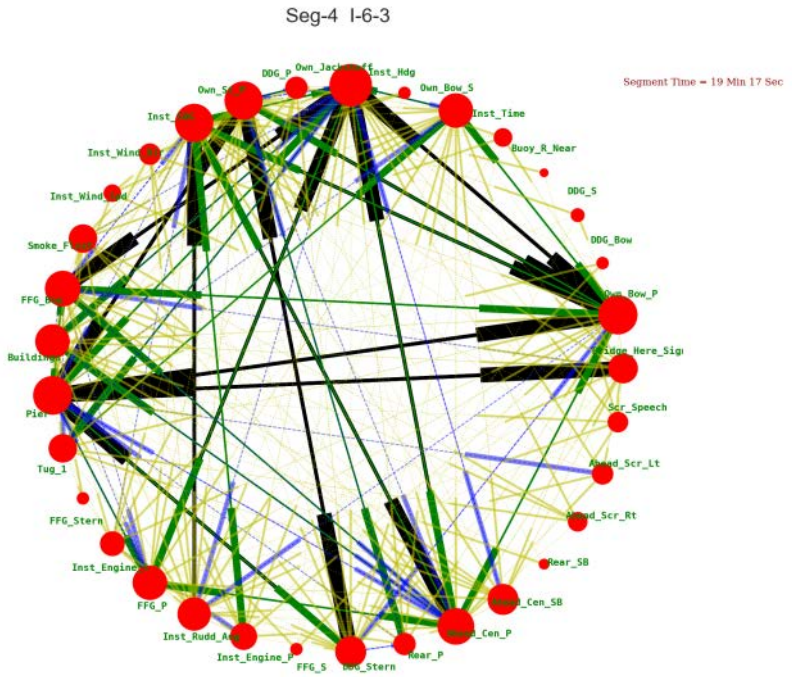


Figure 85. I-6-3 Eye Scan Network Graph in Segment 4

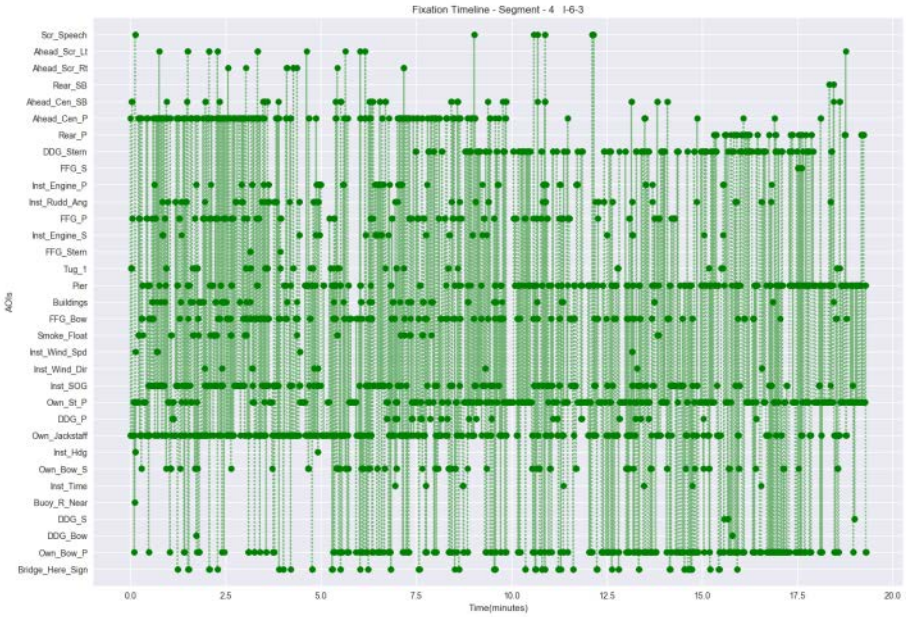


Figure 86. I-6-3 Eye Scan Timeline Graph in Segment 4

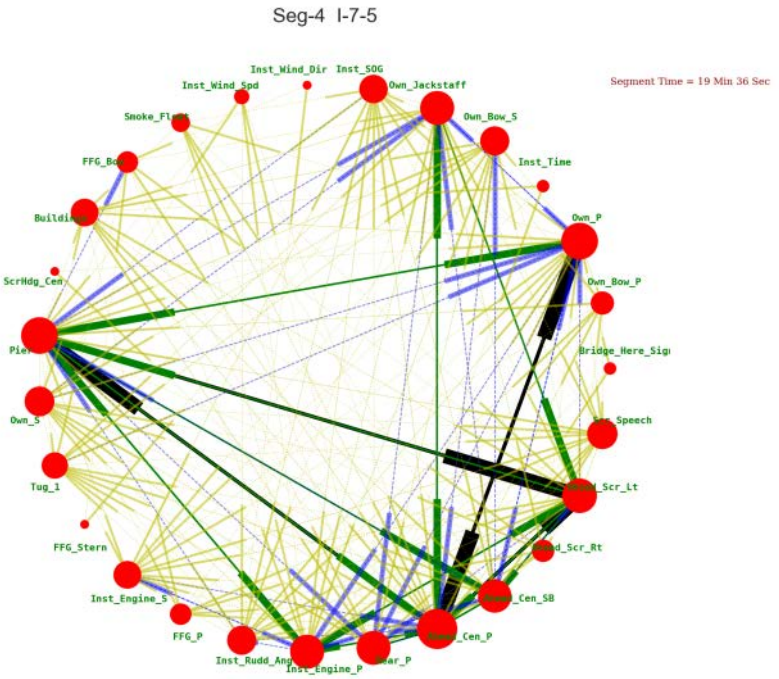


Figure 87. I-7-5 Eye Scan Network Graph in Segment 4

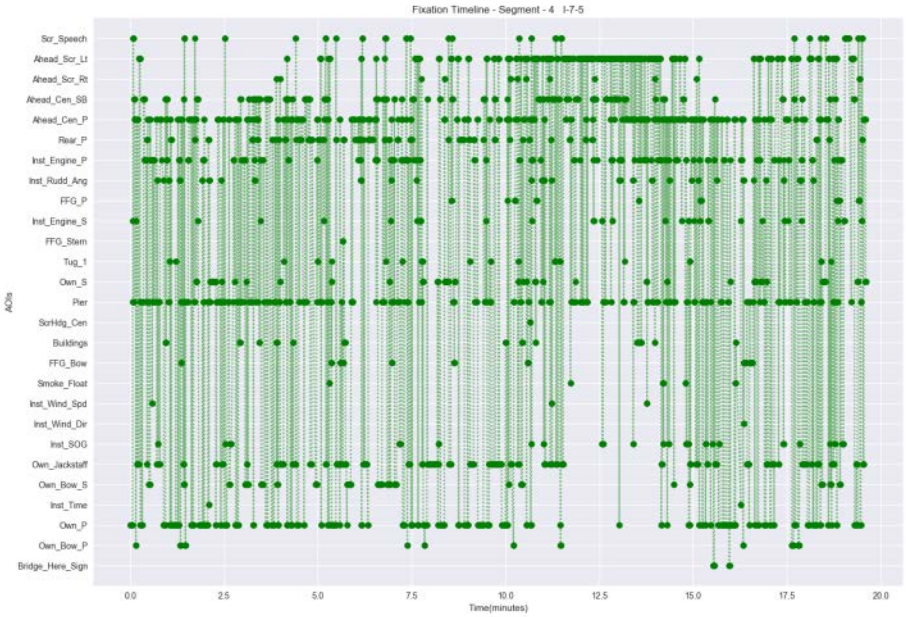


Figure 88. I-7-5 Eye Scan Timeline Graph in Segment 4

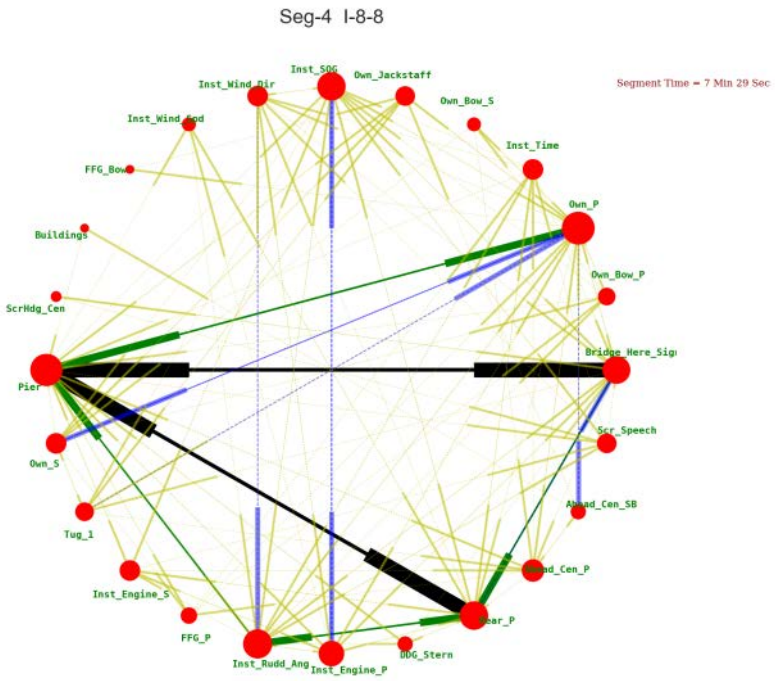


Figure 89. I-8-8 Eye Scan Network Graph in Segment 4

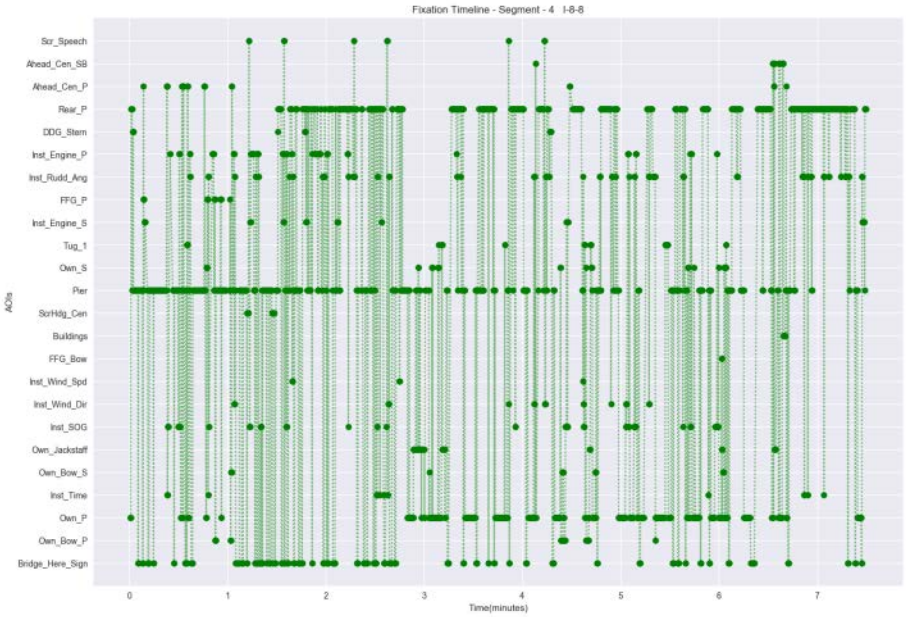


Figure 90. I-8-8 Eye Scan Timeline Graph in Segment 4

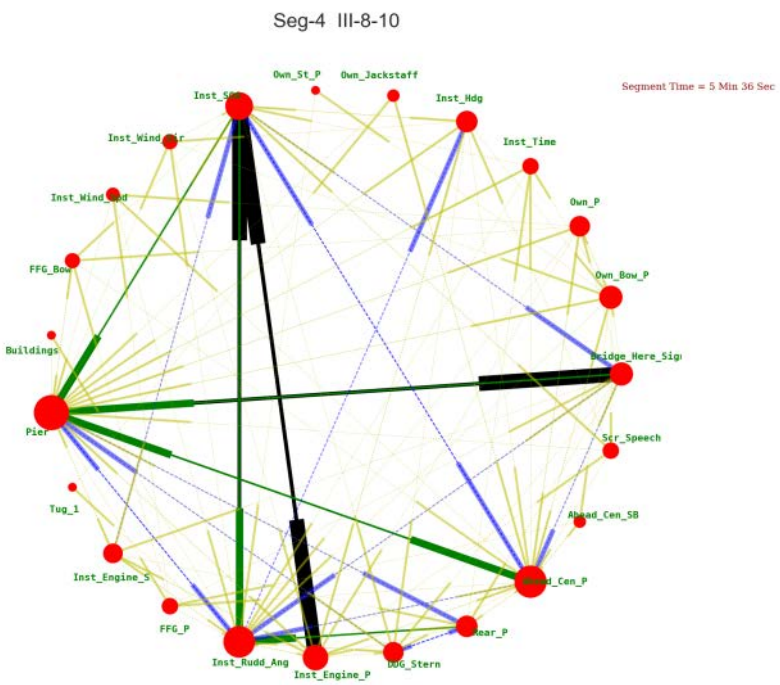


Figure 91. III-8-10 Eye Scan Network Graph in Segment 4

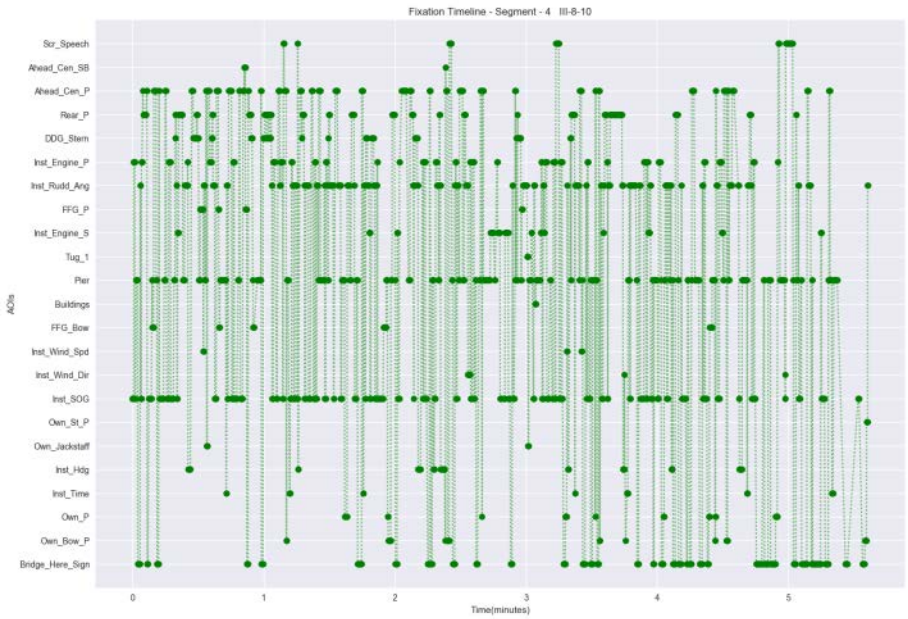


Figure 92. III-8-10 Eye Scan Timeline Graph in Segment 4

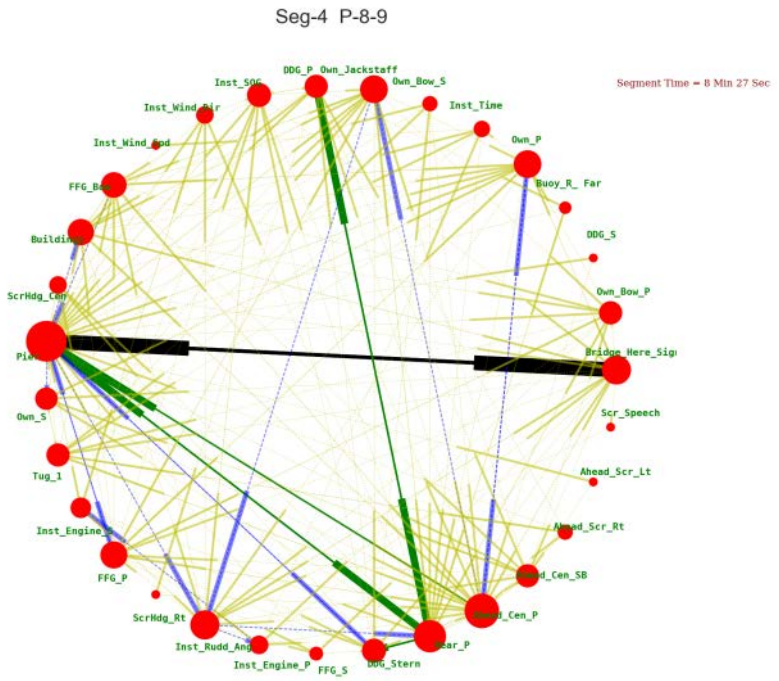


Figure 93. P-8-9 Eye Scan Network Graph in Segment 4

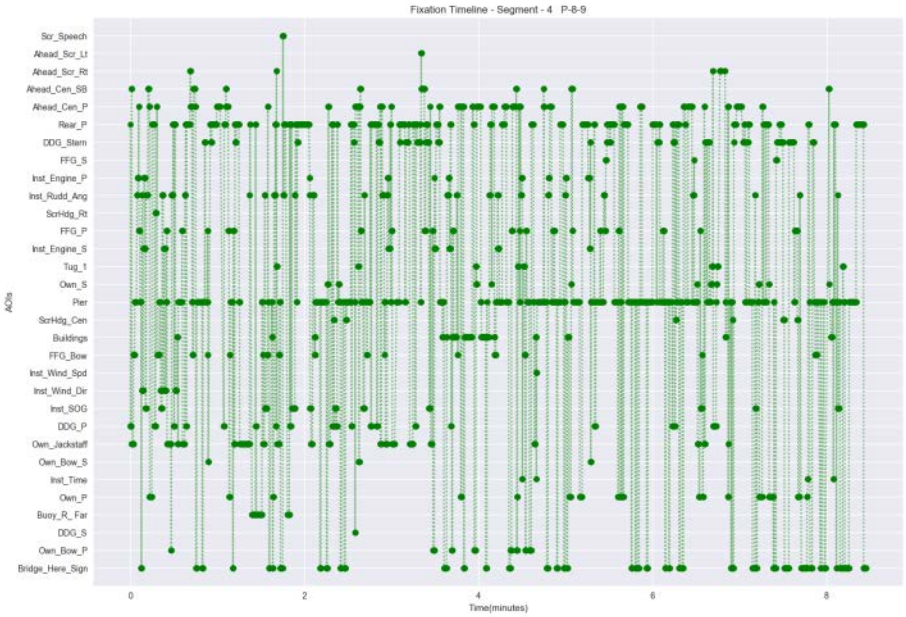


Figure 94. P-8-9 Eye Scan Timeline Graph in Segment 4

LIST OF REFERENCES

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological Review, 111*(4), 1036–1061.

Bratt, E. O., Schultz, K., & Peters, S. (2007). Challenges in interpreting spoken military commands and tutoring sessions responses. *Grammar Engineering Across Frameworks 2007,* Stanford, CA.

Cohn, J., & Patrey, J. (2001). Virtual environments as a multi-modal "real world" laboratory for training. In *Human Factors and Ergonomics Society Annual Meeting Proceedings* (1891–1895).

Davis, P. K. (1995). Distributed interactive simulation in the evolution of DOD warfare modeling and simulation. *Proceedings of the IEEE, 83*(8), 1138–1155.

Falkmer, T. (2001). Eye movement recording as a tool for accident in-depth investigations: A pilot study. In *Proceedings of the Conference Road Safety on Three Continents.* Pretoria, South Africa, 20–22 September 2000 (VTI KONFERENS 15A, No. 15A).

Forsman, F., Sjors, A., Dahlman, J., Falkmer, T., & Lee, H. C. (2012). Eye tracking during high speed navigation at sea. *Journal of Transportation Technologies, 2*(3), 277.

Grassi, C. R. (2000). *A task analysis of pier-side ship-handling for virtual environment ship-handling simulator scenario development* (Master's thesis, Naval Postgraduate School). Retrieved from https://library.nps.edu/nps-theses

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford: Oxford University Press.

Horiguchi, Y., Sawaragi, T., Nakanishi, H., Nakamura, T., Takimoto, T., & Nishimoto, H. (2015). Comparison of Train Drivers' Eye-Gaze Movement Patterns Using Sequence Alignment. *SICE Journal of Control, Measurement, and System Integration, 8*(2), 114–121.

Horiguchi, Y., Suzuki, T., Sawaragi, T., Nakanishi, H., & Takimoto, T. (2016). Extracting train driver's eye-gaze patterns using graph clustering. *IFAC-PapersOnLine, 49*(19), 621–626.

Iyengar, J. V., Hargett, G. H., & Hargett, J. D. (1999). Military development and applications of simulation systems. *Journal of International Information Management, 8*(2), 5.

Karsarkis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. (2001). Comparison of expert and novice scan behaviors during VFR flight. *Proceedings of 11th International Symposium on Aviation Psychology*, pp. 1–6.

Keller-McNulty, S., Bellman, K. L., Carley, K. M., Davis, P. K., Ivanetich, R., & Laskey, K. B. (2006). *Defense modeling, simulation, and analysis: Meeting the challenge.* Washington, DC: National Research Council, The National Academies Press.

Kennedy, Q., Fricker, R., Appleget, J., Alt, J., Patton, S., Whitaker, L., Nesbitt, P., Yang, J.H., & Huston, J. (2013). *Understanding optimal decision making in war-gaming*. Monterey, CA: TRADOC Analysis Center.

Kincaid, J. P., & Westerlund, K. K. (2009, December). Simulation in education and training. In *Winter Simulation Conference*, pp. 273–280.

Kirby, C. E., Kennedy, Q., & Yang, J. H. (2014). Helicopter pilot scan techniques during low-altitude high-speed flight. *Aviation Space and Environmental Medicine, 85*(7), 1–5.

Kirschenbaum, S. S., Wong, J. H., & Peters, S. (2010, September). Tutoring Ship Handling Skills: Intelligent Tutor and Expert Model in a Simulation. In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting.* (Vol. 54, No. 27, pp. 2279–2283). Los Angeles, CA: SAGE Publications.

Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Human Factors, 49*(3), 347–357.

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50*(6) 864–878.

Sullivan, J., Yang, J. H., Day, M., Kennedy, Q. (2011). Training simulation for helicopter navigation by characterizing visual scan patterns. *Aviation Space and Environmental Medicine, 82*(9), 1–8.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research, 22*(22), 4673–4680.

Van Dongen, S. M. (2001). *Graph clustering by flow simulation* (Doctoral dissertation). Dutch National Research Institute for Mathematics and Computer Science, University of Utrecht, Utrecht, Netherlands.

Van Dongen, S. (2008). Graph clustering via a discrete uncoupling process. *SIAM Journal on Matrix Analysis and Applications, 30*(1), 121–141.

Wong, J. H., Kirschenbaum, S. S., & Peters, S. (2010). *Developing a cognitive model of expert performance for ship navigation maneuvers in an intelligent tutoring system.* Naval Undersea Warfare Center Division, Newport, RI.

Yang, J. H., Kennedy, Q., Sullivan, J., & Fricker, R. D., Jr. (2013). Pilot performance: Assessing how scan patterns & navigational assessments vary by flight expertise. *Aviation Space and Environmental Medicine, 84*(2), 116–24.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California